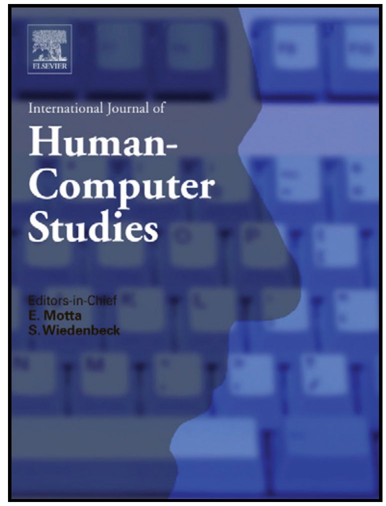


www.elsevier.com/locate/ijhcs

# Supporting intelligent and trustworthy maritime path planning decisions

M.L. Cummings, Mariela Buchin, Geoffrey Carrigan, Birsen Donmez

PII: S1071-5819(10)00064-9
DOI: doi:10.1016/j.ijhcs.2010.05.002
Reference: YIJHC 1592

To appear in: *Int. J. Human–Computer Studies*

Received date: 4 July 2009
Revised date: 14 April 2010
Accepted date: 12 May 2010

Cite this article as: M.L. Cummings, Mariela Buchin, Geoffrey Carrigan and Birsen Donmez, Supporting intelligent and trustworthy maritime path planning decisions, *Int. J. Human–Computer Studies*, doi:10.1016/j.ijhcs.2010.05.002

# Supporting Intelligent and Trustworthy Maritime Path Planning Decisions

M.L. Cummings*
Mariela Buchin
Geoffrey Carrigan
Birsen Donmez

Humans and Automation Laboratory
Massachusetts Institute of Technology
77 Massachusetts Avenue, 33-311
Cambridge, MA 02139
(617) 252-1512
Missyc@mit.edu
mariela@alum.mit.edu
Geoffrey.Carrigan@gmail.com
bdonmez@mit.edu



**Keywords: path planning, decision support, trust, maritime, navigation**

* Corresponding author

**Abstract**

The risk of maritime collisions and groundings has dramatically increased in the past five years despite technological advancements such as GPS-based navigation tools and electronic charts which may add to, instead of reduce, workload. We propose that an automated path planning tool for littoral navigation can reduce workload and improve overall system efficiency, particularly under time pressure. To this end, a Maritime Automated Path Planner (MAPP) was developed, incorporating information requirements developed from a cognitive task analysis, with special emphasis on designing for trust. Human-in-the-loop experimental results showed that MAPP was successful in reducing the time required to generate an optimized path, as well as reducing path lengths. The results also showed that while users gave the tool high acceptance ratings, they rated the MAPP as average for trust, which we propose is the appropriate level of trust for such a system.



## 1.0 Introduction

After a significant worldwide decline in serious navigation-related commercial maritime accidents from 1987-2002, the past five years have seen a significant spike in these accidents to levels not seen in more than 20 years (Richardsen, 2008). This recent trend is also reflected in United States maritime operations with a recent similar spike in US Navy accidents. Furthermore, in the past 25 years, the National Transportation Safety Board (NTSB) has investigated more than 50 collisions (with other ships and infrastructure such as bridges), and running aground incidents[†]. Collisions and groundings now account for 60% of the most costly maritime accidents, and in the current climate, a ship is twice as likely to be involved in a serious

[†] The NTSB does not investigate every incident, only those of major significance.

grounding and collision as compared to only five years ago (Richardsen, 2008). In all cases, human error is cited as a central cause, but other more latent causes that have been identified include a lack of situation awareness, an undersupply of crew, and high workload in navigation settings.

In coastal and high density traffic settings, when unexpected events occur that require immediate route replanning, such as erratic movements of other maritime traffic, resultant plotting and charting can take several minutes, even with electronic displays. Navigation in congested and littoral regions causes significant navigator stress (Grabowski & Sanborn, 2003), as course replans and small adjustments occur frequently, increasing navigator workload. Increases in mental workload, shown to be intricately linked with losses of situation awareness (Endsley, 1993), can lead to increased chances of allisions or collision (Grabowski & Sanborn, 2003).

Navigation is an inherently complex cognitive task since it typically involves multiple variables, many of which are uncertain (such as currents and other ships' movements) that must be optimized to some objective function, often under time pressure (Hutchins, 1995). Moreover, navigation in coastal and especially harbor areas is especially demanding and in military settings, can require up to ten different people: the navigator, assistant to the navigator, navigation plotter, navigation bearing recorder/timer, starboard and port pelorus (a compass attached to a sighting telescope) operators, restricted maneuvering helmsman, quartermaster of the watch, restricted maneuvering helmsman in after-steering, and fathometer (depth) operator (Hutchins, 1995). Planning courses under time pressure, while not typically an issue for open ocean vessels, is particularly problematic for military littoral warships and fast patrol boats.

For ships equipped with the most modern technology (typically large commercial vessels), a merchant ship navigator can plot a course on an electronic map with zoom capability, which can be configured to show different layers of information such as weather and depths. In addition, some ships have radar systems that automatically identify and track other vessels in the water, such as the Automatic Identification System (AIS) which can transmit positions and speeds to an electronic display, if a ship has that capability. However, there is currently a lack of sufficient integration between the systems (Lee & Sanquist, 1996; Perrow, 1984), creating more demand on operators to process and integrate the data presented to them (Lee & Sanquist, 2000; Urbanski, Morgas, & Czaplewski, 2008). Moreover, such electronic aids have been shown to be useful in low stress settings, but problematic in high stress scenarios (Grabowski & Sanborn, 2003). This problem is not just a maritime one, as the aviation industry has struggled with similar issues of increased workload with increased automation (Billings, 1997).

Not all maritime organizations use these electronic tools, and many ships, including most US military ships, still rely on the traditional paper chart method for navigation. The tools used in plotting ships' path can include an alidade, which is a device that sights a landmark to measure the spatial relationship between the home ship and that landmark, the hoey, which is a one-arm protractor used in translating the angular relationship between the home ship and a landmark into a map bearing, parallel rulers, parallel motion protractors, compasses, distance scales and dividers for measuring distances (Hutchins, 1995). These devices all have degrees of error in accuracy, and training and experience play a significant role in path quality and time to plot a path.

Time to plot a path can be a significant stressor in high workload navigation environments such as dense coastal settings. Personnel who plot courses on paper charts experience high mental workload when faced with the need to rapidly replan and chart in the face of new information, such as the presence of unexpected radar contacts or rapidly advancing weather. In some military operations, some ship captains will bring their vessels to a halt while attempting to replan a new course because of an unexpected event, which has clear negative mission implications, particularly in terms of time pressure.

We propose that both in paper and electronic chart systems, what is needed to reduce workload in time-pressured navigation tasks is a decision support tool that integrates the various sources of critical navigation information via an automated path planner and a user-centered visualization. Leveraging an intelligent path planning tool could greatly increase the accuracy and speed of planning a path, as well as reduce workload and error, and possibly manning requirements. While current electronic displays provide descriptive representations of the navigation environment and some limited predictions (e.g., where contacts are likely going), no tool currently in operational settings has effectively leveraged some form of intelligent decision support to aid humans in this demanding task.

Little research has investigated the use of automated path planning in maritime navigation. Rothgeb (2008) demonstrated that a fuzzy logic neural net could be used to identify high risk areas of transit given known contacts, as well as generate a recommended course based on safe areas. However, this research was focused on contact management, and not on the more holistic problem of path planning given additional variables such as weather and operator experience. In another related effort, Smierzchalski, et al. (1998), developed an automatic path planner that accounts for surrounding contacts and their future positions, as well as physical characteristics of the ship such as weight, center of gravity, and size of control surfaces. Their proposed algorithm, EP/N++, a variant of the EP/N (Evolutionary Planner/Navigator) algorithm for mobile robots (Xiao, Michalewicz, & Trojanowski, 1997), randomly generates acceptable paths for getting a ship from one point to another as a function of least cost. This randomized approach causes the solutions to be near-optimal at best, with the optimal solution traded for algorithm speed. This research is somewhat limited, as the proposed algorithm only takes into account up to three contacts in the vessel's area of observation, and it does not address uncertainties for future contact positions. In addition, while the algorithms were tested in limited scenarios, no human-in-the-loop trials were ever conducted with any functional decision support tool based on the automated path planner.

Although automated path planning research in maritime navigation is limited, there is extensive research in the field of robotic path planning which can provide useful insights to maritime navigation. Path planning in navigation is a large area of research in the computer science field (Winston., 1992), with significant research conducted in robotic path planning (e.g., (LaValle, 2006; Russell & Norvig, 2003; Thrun, Burgard, & Fox, 2005). As will be discussed in more depth in the next section, given this previous research, we elected to use the A* algorithm for our automated path planner, which is an informed search method that can quickly find an optimal path to a destination, given our relatively constrained state space.

While an automated path-planner algorithm that is accurate and fast is critical for the maritime navigation problem, equally as important is the development of an operator decision support tool that maintains high operator performance, while also reducing mental workload. Users need to understand the limitations of such automated planning tools in order to know when they are correct (Layton, Smith, & McCoy, 1994). This issue raises another critical design consideration in the development of a maritime path planning tool, which is trust. New technologies in complex systems such as automated path planners in maritime settings face the challenge of gaining an acceptable level of trust from the operator before the system is accepted. When an operator has too little or too much trust in a system, the system has the potential to be dangerous. Distrust may lead to system disuse, and over-trust may lead to inappropriate reliance on a system (Parasuraman & Riley, 1997).

Trust in intelligent decision support tools is affected by the reliability of the automation (Lee & Moray, 1992; Parasuraman & Riley, 1997). Research has shown that when automation reliability is in doubt, users' trust in the automation significantly drops, causing more reliance in manual methods (Ruff, Narayanan, & Draper, 2002), which then negates the usefulness of the automation. Moreover, the perceived reliability that a user attributes to automation is often related to how the information from the automation is conveyed to the user (Parasuraman & Riley, 1997). Increasing system uncertainty has also been shown to be a source of distrust for operators of automated systems (Uggirala, Gramopadhye, Melloy, & Toler, 2004). Uncertainty can stem from the environment, but it can also come from automated sensing and computation, so when designing an automated path planner, a designer needs to consider the impact of uncertainty from both of these sources.



In summary, there is a clear need for more effective navigation decision support in maritime settings, particularly in coastal settings. While a few researchers have examined how different intelligent algorithms could be applied to limited aspects of this problem, no previous research has looked at the intersection of human-algorithm performance for the global maritime path planning task. Moreover, given the importance of user trust and acceptance in successful transitions of such technologies, any path planning decision support tool should be explicitly designed to mitigate uncertainty and enable user understanding of automation-generated solutions. To this end, in the next section we discuss the development of a maritime path planning decision support tool, primarily targeting coastal and high density traffic settings, that allows operators the ability to leverage automation to quickly generate multiple path options that account for contacts, weather, and depth restrictions, as well as allowing them the ability to adjust their level of risk.

## 2.0 Decision Support Design

Since coastal navigation (in and around land, harbors, and shipping channels) represents the most demanding phase of maritime navigation, it is important to determine the baseline processes and tasks involved in safe navigation in order to develop a comprehensive and functional decision support tool. In order to determine the display requirements for such a decision support tool, a cognitive task analysis (CTA) was conducted (Schraagen, Chipman, & Shalin, 2000), which yielded information about the knowledge, thought processes, and goal structures of the maritime navigation process. Our CTA included interviews with U.S. Navy personnel (5) and Northeast Maritime Institute personnel (4), all with significant experience in maritime navigation (~15 years on average). The end result was a list of information requirements, seen in Table 1 (the detailed CTA can be found in Carrigan (2009)). In order to facilitate understanding of the decision support display that stemmed from these requirements, the following section details the general workflow in maritime settings, which was also derived from the CTA.



### 2.1 Maritime Navigation Workflow

Generally while navigating in real time, if an unexpected event occurs such as lowered visibility or a ship finds itself on a collision course with another vessel, the navigator must determine acceptable alternate routes, choose and plot this route on the map, measure the route's orientation with a compass to determine heading, and also determine path length which is then translated into time. Depending on the situation, multiple iterations of re-planning may be required to find a suitable solution. The tools needed for conventional paper-based plotting include a bulky paper chart, tracing paper (so to not permanently mark expensive charts), a parallel plotter, a compass, and a pencil. These re-planning events are laborious and time-consuming and even under the best of circumstances with a skilled navigator, this task can take between 3-5 minutes to complete. Even small course changes take about a minute to create.

**Table 1: Information Requirements**

| Type | Requirement Description |
|---|---|
| Environment Information | Current speed and heading |
| | Current and expected depth along projected path |
| | Current and expected visibility along projected path |
| | Geo-spatial boundaries of operating area |
| | Visual navigation lanes |
| | Hazardous/restricted areas |
| | Planned course, highlighted if blocked |
| | Visual indication of allowable paths |
| | Start and final destination or goal location on map |
| | Areas where collision is possible or uncertain with obstacles (e.g., shoals, reefs, etc.) |
| | Ability to compare different routes |
| Contact Information | Geo-spatial location of all surrounding contacts |
| | Each contact bearing, speed, and whether course is opening/closing |
| | Contact path: past, present and future |
| | Contact location on path |
| | Marking to distinguish contacts with Automatic Identification System (AIS) data |
| | When and where ownship is on a projected collision course with a contact |
| | When path is inaccessible, explain why |



In addition to the need to find a course that avoids obstacles in coastal regions (such as shoals), navigators must consider the local traffic. Contacts can be acquired through sonar, radar, visual, or the AIS electronic data described earlier. Other environmental concerns in maritime navigation settings include currents, especially as they relate to shallow areas, and weather, particularly visibility because navigators rely on seeing the lights from other ships as well as buoys and other lane markers to ensure that they are in the proper channels, and clear of obstacles.

Given this workflow and the detailed requirements as outlined in Table 1, an automated path planner was developed with the primary purpose of reducing the workload of the navigator or any higher-level decision maker such as the ship's commander in a time-pressured setting. By reducing workload, we explicitly mean a reduction in time to plan (which is a primary task performance-based measure) with resultant path quality that 1) does not violate any obstacles, and 2) the new path is at least equal to or shorter than a path generated from a paper chart. As will be described in more detail in the next section, the automated path planning tool incorporates knowledge of obstacles and contacts to provide a path that is not only efficient, but also safe.



## 2.2 Maritime Automated Path Planner (MAPP)

Given the information and display requirements discussed previously, a prototype display was developed to aid maritime personnel rapidly plan paths (Figure 1). This navigation planning display, called MAPP (Maritime Automated Path Planner), embeds both manual and automated features that allow users to plan and replan paths quickly, with automated constraint checking discussed in detail below. MAPP was originally designed for a handheld computer since the need for mobility was expressed by the subject matter experts. In order to support rapid replanning, no information was layered and direct manipulation was supported through touch and/or a stylus.

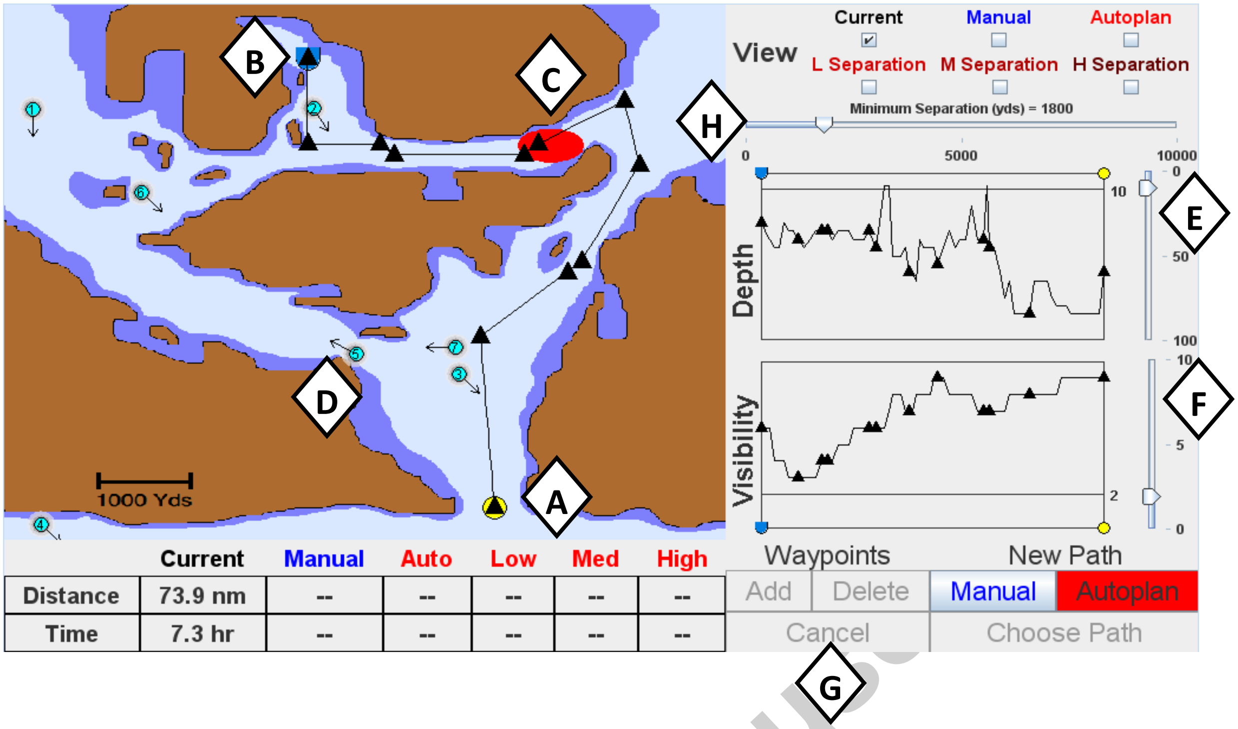


**Figure 1: Maritime Automated Path Planner (MAPP)**

As seen in Figure 1, the primary display is a map representation of the area of operation. Diamonds A and B denote the start and end points of the example, connected via the current path in Figure 1, and C marks a danger area that represents some time-critical hazard area to be avoided such as an oil spill or security threat. The small circles, for example as marked by diamond D, represent contacts and are overlaid with trajectory vectors that represent both speed and heading. Such notations are common in air traffic control settings and are helpful for projecting possible future conflicts. The trend graphs on the right show the expected depths and visibilities that the ship will experience along the currently selected path, which provide visual representations of what sections of the current path will come close to pre-specified threshold levels that can be set via sliders in diamonds E and F.

In MAPP, users can either ask the automation to plan or replan a path, or they can elect to plot a path themselves (the Manual button in the lower right of Figure 1). The "Autoplan" button can generate multiple obstacle-free paths between the ship's current position and the goal by taking user input about what contact separations, depths and visibilities are deemed acceptable. The “Manual” tools allow users to create their own paths by adding, deleting, and dragging waypoints in the map (diamond G), with some constraint checking concerning acceptable depths and visibilities. Both modes are discussed in more detail below.

The Manual mode is important because not all maps are correct, both in their representation of the world and where users think potential problems might be. The Manual mode allows users to avoid paths that may go through obstacles that are not documented in paper or electronic charts. While called the Manual mode, this mode actually is quasi-automated in that users specify waypoints they would like to add, delete, or modify and the automation generates path and heading markings between the waypoints and checks for obstacles in terms of land, depths, and visibilities. The manual mode gives users the freedom to choose a desired course, while still providing efficient and fast path manipulation, as well as confirmation that the path is safe from hazardous areas and poor weather regions (as determined by the thresholds set by the depth and visibility sliders (E and F diamonds)). In this manual mode, automation is leveraged as a “critic” (Guerlain et al., 1999), so that it can prevent user error, but only for path-related issues (including depth and visibility violations). The critiquing automation in the Manual mode does not take into account any information about the contacts. Incorporation of this information is left to the user in this mode.



The Autoplan mode allows operators the ability to generate several paths based on different criteria. An A* algorithm (Hart, Nilsson, & Raphael, 1968) is embedded in this planning/re-planning advisory tool since it can calculate a fast and optimal solution given an admissible heuristic (Winston, 1992). Moreover, the size of the grid is relatively small, so the amount of memory needed to solve this problem did not warrant using a more complex algorithm, although should the need arise to expand the state space, other algorithms could easily be embedded in this tool. The heuristic is the Euclidean distance from the ship's current position to the goal. Each cell for this tool is 50x50 pixels. Since each node expansion only examines the eight cells adjacent to that node, the resulting path is jagged, which translates to multiple heading changes and is unacceptable in maritime settings. To rectify this, we used a modified version of the Jacobi relaxation algorithm to smooth the paths, such that a point in path is the averaged value of the previous and next points in the path (Goldstine, Murray, & von Neumann, 1959). Every new value is also checked against the obstacle database to ensure that the new point is safe.

While accounting for depth obstacles, the algorithm also considers visibility, a critical weather consideration in coastal navigation. Since weather is highly uncertain, a simple model was developed that converts forecasted visibility from a simulated weather chart to cell upper and lower bounds, i.e., some regions of the operational area are more likely to experience lower visibility than others which is typical in actual maritime settings. In terms of the A* algorithm, a cell is classified as a visibility obstacle if the cell's visibility lower bound is less than the cutoff visibility value specified by the user in slider F (Figure 1).

In addition to depth and visibility, the algorithm also accounted for possible contact proximity. Since an automated path planner is more useful if it can predict where other contacts could be in the future, every contact's likely future trajectory was modeled through dead-reckoning, which means a future path was projected from a contact's current position based on last known speed and heading. While this process can carry significant uncertainty, recent technologies such as AIS have significantly reduced this variability. However, since not every ship has such advanced technology, a circular area was assigned to each contact on the interface (as well as coded into the algorithm), which conveyed increasing probabilistic uncertainties for greater radial distances from the contact's last known position.

One important result from the CTA was the need to support different experience levels (i.e., novices versus experts) in the path planning process. The manifestation of this finding is the display implementation of a user's ability to accept greater risk associated with passing ships/contacts. To support this experience-based need, three display features were added. The first was a minimum separation slider (diamond H) that allows users to adjust the absolute minimum separation when passing a contact, which is then considered by the algorithm. Second, three rings are shown around each contact so that users can estimate the level of risk they are willing to assume when transiting near specific contacts. These three grey-shaded rings correspond to three levels of desired separation (low, medium, and high), with dark representing low separation areas and light grey representing high separation areas. As the separation slider is adjusted, these rings adjust accordingly.

The last feature included to support different user risk profiles is a what-if tool in the upper right of Figure 1, which allows users to select up to six different paths to view, which include the current path, a Manual path, an Autoplan path that does not consider contacts, and/or L(ow), M(edium), and H(igh) Separation paths for contact avoidance. In the manual mode, operators internally account for possible contact behavior, while in the Autoplan mode, the algorithm treats the contacts as obstacles, with the obstacle size driven by the user-specified level of separation. For example, a low separation selection causes the algorithm to consider the contact's size to be the minimum distance set in the slider, and it only considers that contact as an obstacle for the intersection of that contact's dead reckoning path and the A* proposed path.

The last interface element of note in Figure 1 is Table E in the lower portion that provides summary data for the different paths that can be selected. This table presents quantitative information about each path in terms of length in nautical miles, and how long it will take the vessel to traverse the entire path at the current ship's speed. Figure 2 represents what MAPP looks like when three of the six possible paths are selected for comparison. Users can see both proposed automated and manual paths, as well as vertical path profiles in comparison to the current path. Similar approaches to information integration in maritime displays have been shown to improve navigation performance (Sauer et al., 2002), and even enable navigation by a single operator in some cases (Schuffel, Boer, & van Breda, 1988).



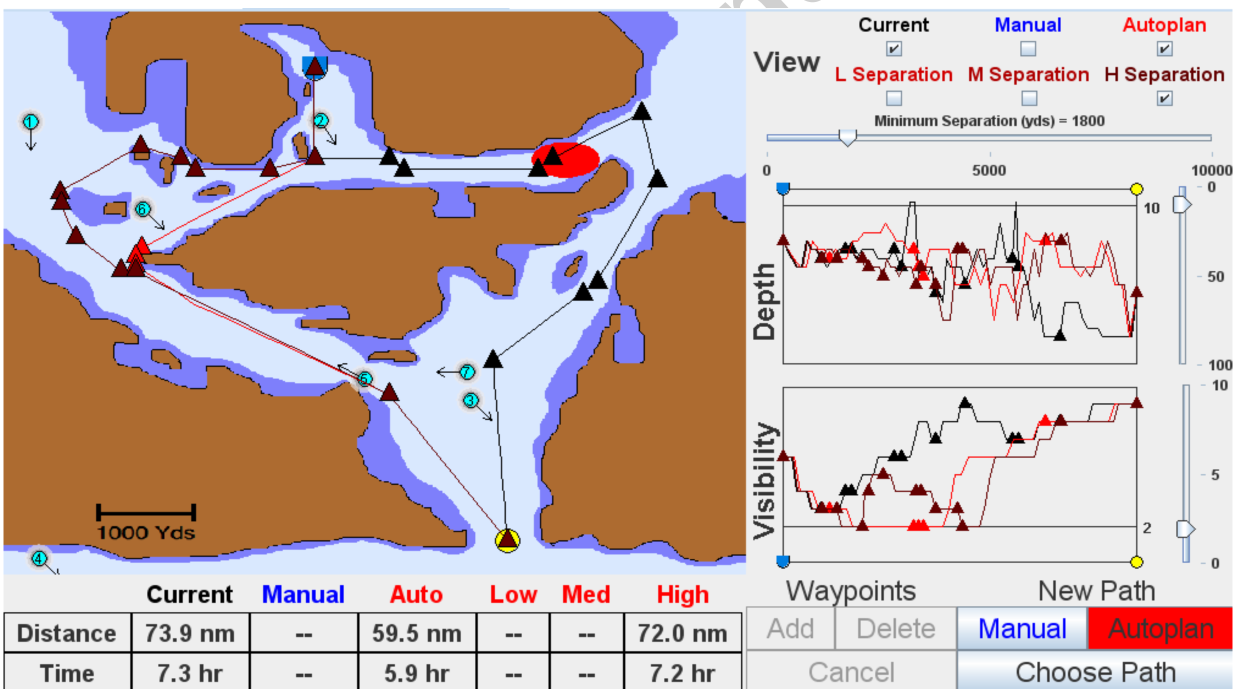


**Figure 2: Leveraging What-If Capabilities in MAPP**

## 2.3 Designing for Trust

While there are a number of specific trust design criteria that could be considered in the development of an automated decision support tool (see Lee and See (2004) for a review), for this maritime path planning problem, we identified two significant trust issues through interviews with subject matter experts: 1) Consideration and representation of environmental uncertainty in the automation's solutions, and 2) Cultural biases that could cause possible misuse or disuse of the path planner. These are discussed in turn.

Environmental uncertainties that cause the greatest stress for navigators include future contact positions and unexpected changes in the physical environment such as weather and hazard areas, which cause the need for rapid replanning. We addressed this concern about uncertainty representation in two ways in MAPP. First, we chose to embed uncertainty computationally through the A* algorithm cell costs, weighted per user specifications, i.e., the variables in the slider bars and the low, medium, and high separation check boxes. For example, given weather conditions, depth constraints, and established contacts, users can set boundaries (e.g., do not go into shallow areas) that then can be considered by the automation in finding the most optimal path. Such an approach allows users to tailor the automation to fit their own perception of environmental uncertainty, e.g., they can make conservative paths with wider margins in the presence of erratic contacts. Since users determine the constraints for the automated solutions, the automation is more transparent which is a critical consideration in increasing user trust (Sheridan, 1988). A number of aviation accidents have been caused by opacity in displayed automation solutions, particularly in navigation settings (Billings, 1997), so transparency is a critical design consideration.

Secondly, we addressed the uncertainty problem though direct visual representation in the form of contact position shaded concentric circles. Previous research has shown that when uncertainty exists in automated solutions, this uncertainty should be conveyed in order for users to develop appropriate trust in the automated system (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003; Lee & Moray, 1992; Lee & See, 2004). In MAPP, the concentric circles, representing low, medium, and high areas of risk, grow as uncertainty in contact position increases. This technique leverages direct perception (Gibson, 1979) in that users can immediately see how far a proposed path is from a potential contact. Similar techniques have been used in other trust studies which show that when an interface explicitly depicts information source reliability, users develop appropriate levels of trust (MacMillan, Entin, & Serfaty, 1994; Montgomery & Sorkin, 1996).

In addition to the concerns about the incorporation of environmental uncertainties, interviews also highlighted potential cultural biases that could negatively influence trust, which is a well-established trust in automation problem (Lee and See, 2004). Interviews indicated that the civilians appeared to be more receptive of proposed automated planning tools than did military personnel, who expressed significant reluctance. In order to address these possible user acceptance issues, which could lead to operator misuse or disuse, we incorporated what-if tools in the interface through the ability to select different paths based on different subjective levels of risk. In MAPP, users have the ability make comparisons both within different risk levels for automation-generated paths, as well as the ability to compare their own manual paths to the automation-generated paths. Such use of what-if sensitivity analysis tools can increase transparency and understandability of automation-generated solutions, which again are key in increasing overall user trust (Lee & See, 2004; Sheridan, 1988).

## 3.0 Methods

Regardless of whether users prefer the more manual or automated modes in maritime path planning settings, we propose that MAPP should significantly reduce the workload of operators in terms of path quality and speed of path planning for maritime applications. MAPP automates the tedious work of actual plotting paths, while ensuring that routes do not violate depth, visibility, and obstacle constraints, which can be tailored by users. The ability for users to explicitly set constraints within the automation, as well as tailor solutions to different risk levels was hypothesized to promote greater than average operator understanding and perceived automation transparency, which can lead to greater trust. Lastly, interviews suggested that military personnel may distrust automated path planners more than their civilian counterparts, so this and the previous hypotheses were tested in an experiment, described in the next sections.

### 3.1 Participants

Eight military personnel, both U.S. Coast Guard and Navy, and eight civilians from a local sailing pavilion and the Northeast Maritime Institute with significant professional experience in navigating in coastal environments were recruited for this experiment. The average age of the military personnel was 31 years, with a standard deviation (sd) of 7 years. The average age of the civilians was 41 years (sd=10). Average years of professional navigation experience were 15 (sd=11), with civilians averaging 22 years and military personnel averaging 8 years. All participants were male.

### 3.2 Experimental Apparatus and Procedure

All participants in this experiment used both the traditional paper and pencil method and MAPP to plan paths. The paper charts replicated the digitized map used for MAPP, including depth markings, and were 12in x 12in sheets of paper, typical of actual maritime paper charts.

The paper charts were covered with clear tracing paper, as is normally done in the field to prevent marking expensive charts. Corresponding weather charts for the map were also given to participants on a 12in x 12in sheet of paper to simulate typical weather forecasting data. Standard plotting tools were given to participants, which included a ruler, parallel ruler, compass, slide ruler, and pencil.

The intelligent maritime path planning tool, as seen in Figures 1 and 2, was loaded on a Sony Vaio UX390N Micro PC, with a 4.5in screen and a resolution of 1024x600. Users could interact with the device via a touch screen or a pointing stylus. Given the portable nature of MAPP, all experiments were held in the location where participants worked. After signing consent forms, participants were shown a slide tutorial to familiarize them with MAPP, and then were given 15 minutes to practice using it, as well as the paper charts. Participants were encouraged to ask the observer questions during the tutorial and practice session as well as during the actual experiment if necessary.

After the tutorial and practice sessions were completed, each participant then completed two test scenarios. For each scenario, participants were given a map with a current path already planned between a vessel's current position and its goal. An obstacle, such as an oil spill or restricted water notice, was presented that required the participant to create a new path to the goal. Participants could generate one or more paths for comparison, but were required to submit a single final path. With MAPP, generating paths could be accomplished using the manual path planning function or the Autoplan function, which accounted for contacts, weather, and depth as described previously. With the paper charts, participants could use the depth and weather charts, and were also provided a contact information sheet which listed the contacts and their last known positions and headings. Such lists are common in actual maritime settings. After the test scenarios, each participant was asked to complete a questionnaire assessing trust and automation acceptance.

### 3.3 Experimental Design

There were two independent variables for this experiment, the navigation aid (computer or paper) which was repeated across participants with the order counterbalanced, and experience (military or civilian). Participants completed two different scenarios, which were counterbalanced across the navigation aid factor to prevent any learning effect. Dependent variables included differential path length, obstacle avoidance, and time to generate a path. Differential path length was the difference in the submitted path length, as compared to the shortest optimal path length. This measure was needed to normalize the path lengths between the two different scenarios.

In coastal and high density settings, time-pressure is a critical consideration since navigators potentially have to continually update their paths and charts. Thus we used time to generate a plan as a primary mental workload measure, i.e., faster plans require less information processing, and thus lower operator utilization. Measuring workload as a function of primary task performance-based measures, particularly speed of performance, is a standard measurement technique, shown to be sensitive to variations in workload (Wierwille & Eggemeier, 1993). The path differential metric is needed since time to plan cannot be considered in isolation without assessment of path quality, in terms of overall human performance.

In addition, trust was assessed via a trust assessment tool developed for human interaction with automated planners in the air traffic control domain (Kelly, Boardman, Goillau, & Jeannot, 2003). We used this scale because of its focus on assessing trust for computer-based planning aids for route generation, but slightly modified it to fit the maritime domain. Additional questions were asked to determine which aspects of the automation influenced participant ratings.

## 4.0 Results and Discussion

Using a 2x2 repeated measures ANOVA ($\alpha=.05$), a statistical difference was found between MAPP and the paper chart with respect to time to plan a path ($F(1,14)=92.47$, $p<.0001$) (Figure 3). There was no significant difference between the civilian and military solutions with respect to time, and there was no interaction effect. In general, time spent using MAPP to replan a path was just over 1.5 minutes, with the average time used for the paper chart for the same scenario at nearly 4 minutes.

With respect to path length, there was again a statistical difference between using MAPP or the paper charts to plot a path ($F(1,14)= 13.21$, $p=.003$). On average, MAPP paths were approximately five nautical miles (nm) shorter than paper routes, about 7% of the total distance. Thus, the shorter paths created by MAPP were more efficient, which has practical implications in that shorter paths lead to both time savings and decreased operating costs. In addition, for this dependent variable, there was a significant difference between the military and civilian solutions ($F(1,14)=18.26$, $p=.001$). There was no interaction effect. As seen in Figure 4, military participants planned longer paths both in paper and in MAPP, which tended to be more conservative near contact paths as compared to their civilian counterparts.

When examining just the paths created in MAPP, while military plans were the most conservative, civilian paths were more conservative than the A* generated paths, in that they were, on average, 1.9 nm (sd=1.8 nm) longer than the "optimal" low separation, highest risk path. In terms of obstacle avoidance, all participants correctly planned their paths, thus verifying their subject matter expertise.

These results should be interpreted in light of the small participant numbers. However, the effect sizes as measured by partial $\eta^2$ were .51 for time to plan and .47 for path differential, indicating medium to strong effect sizes with half of the variance controlled for by the independent variables. Moreover, the participant pool was a highly selective group of actual mariners with significant navigation experience, thus improving overall external validity.

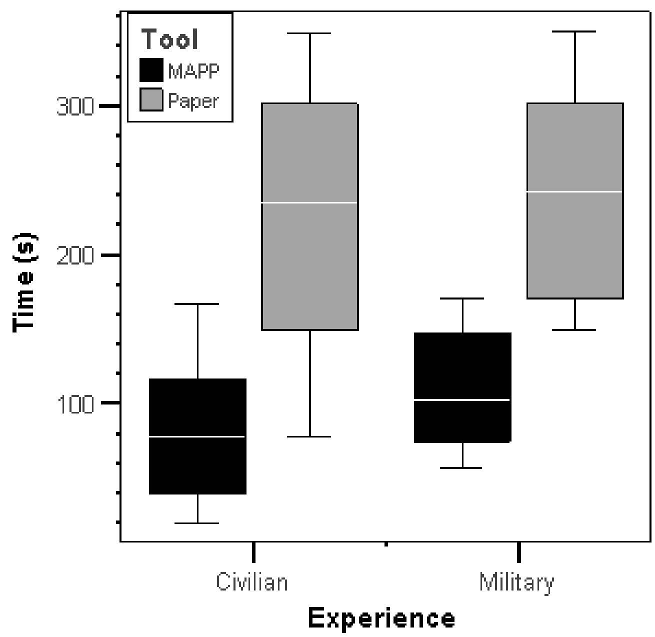


**Figure 3: Path Planning Time**

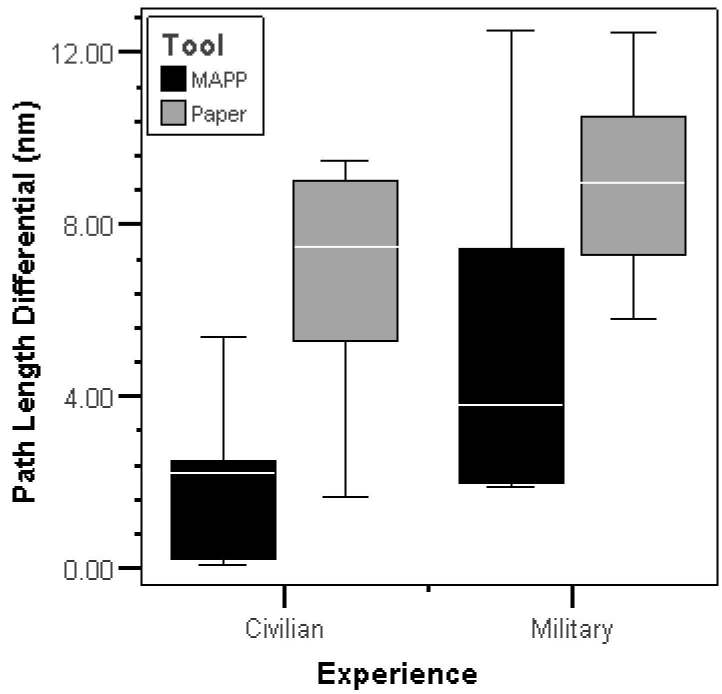


**Figure 4: Differential Path Length**

A regression analysis was conducted to assess if there were any significant predictor variables for high performance. Covariates of age, experience and risk (in terms of path separation) were all checked for significance via a mixed linear model for path length and time, controlling for tool type, background, and their interaction. None of the covariates were found to be significant ($p>.2$). Overall, it appears that the participants' performance was not significantly dependent on age, experience or level of risk. While the lack of significance may be due to the relatively small sample size, it may also indicate that MAPP was simple enough that people were able to use it regardless of past experience.

When examining behavioral interactions with MAPP, in terms of which functionality participants used most (automated or manual), 11 predominantly used the automated path planner (69%), 4 used the manual path planner (25%), and 1 person used both equally. Based on interviews with the SMEs, it was expected that the military users would use the manual functions in MAPP more than the civilians, rather than trusting the automation. This result was seen in that 4 out of 8 military participants used the manual method, while only 1 civilian used the manual function in MAPP. However, this difference was not statistically significant using a chi-squared test.

### 4.1 Trust and Automation Acceptance

Participants were asked a series of questions to determine their overall trust and acceptance of MAPP on five-point Likert scales, depicted in Table 2 where 1 is the lowest and 5 is the highest rating. There were no statistical differences between civilians and military personnel using non-parametric Mann Whitney tests, so the data presented in Table 2 represent participants' overall feedback for these questions. A set of chi square tests comparing all the ratings in Table 2 grouped as either 3 or below (because there were so few ratings of low and fair) and 4 and above showed significant differences for all questions except trust.

In regards to trust, participants were asked, "Please indicate your overall amount of trust." The lowest rating (1) was labeled "No Trust", the middle rating was labeled "Somewhat Trustworthy", and the highest rating was labeled "Complete Trust". Participants could also select intermediate rankings between these. As depicted in Table 2, zero participants selected "No Trust" and only two selected an equivalent fair rating, i.e., between "No Trust" and "Somewhat Trustworthy." The majority of participants (8) found MAPP to be somewhat trustworthy, with one person rating it as completely trustworthy and five choosing a rating equivalent to good in Table 2.

The fact that the majority of participants found MAPP to be somewhat trustworthy, but not completely trustworthy was confirmed with a chi square test that shows no statistical difference in the number of people who felt MAPP should be rated as good to high in terms of trust, as compared to the average and below ratings. When participants were asked to rate how reliable they believed the automation behaved in terms of obstacle avoidance and generating accurate paths, participants found the automation to be above average in reliability ($\chi^2$=9, p=.003). In general, participants also found the interface above average in terms of ease of use ($\chi^2$=6.25, p=.012), with one participant reporting less than average usability. The scores in the automation transparency category, which assessed how well participants understood the behavior of the automation, were the highest of all questions, with a statistically significant number of users reporting good or high understanding of the automation ($\chi^2$=9, p=.003).

**Table 2: Participant Trust and Acceptance Ratings**

| | 1 Low | 2 Fair | 3 Average | 4 Good | 5 High | Mean | Median | $\chi^2$ (1-3, 4-5) |
|---|---|---|---|---|---|---|---|---|
| Overall Trust | 0 | 2 | 8 | 5 | 1 | 3.3 | 3 | 1.0 (p=.317) |
| Obstacle Avoidance Reliability | 0 | 0 | 2 | 10 | 4 | 4.1 | 4 | 9.0 (p=.003) |
| Path Accuracy | 0 | 0 | 1 | 10 | 5 | 4.3 | 4 | 12.3 (p<.001) |
| Automation Transparency | 0 | 0 | 2 | 5 | 9 | 4.4 | 5 | 9.0 (p=.003) |

The results in Table 2 yield two important results. First, the automation transparency question received the overall highest ratings. As discussed previously, transparency was a primary design objective, manifested through direct user access to the variables via the sliders, which allowed users to change the settings to reflect different personal risk profiles. In addition, the embedded levels of risk that users could select in terms of contact avoidance allowed users to tailor the automation to their perceived risk in a comparative fashion to develop what-if scenarios. For those users that elected to use the low, medium, and high separation sliders (N=9), there was a statistically significant relationship between risk taking and trust in that those users who selected high separation paths were more likely to rate the tool higher in terms of trust than those who exhibited riskier behavior, i.e., selecting low separation paths (Spearman $\rho$ =.678, $p$=.045). Another specific design feature included to promote transparency was allowing users the ability to generate and modify automated paths, and compare these to their own manually-generated paths. This behavior was exhibited by 75% of participants.

The second noteworthy result is that despite all the good and high ratings for attributes of the automation (the automation was rated as either good or high 88% across the questions in Table 2), overall participants still only rated MAPP as somewhat trustworthy. While generally embracing the automation for the maritime path planning task and voicing few negative concerns, participants were still relatively lukewarm about trusting the tool, regardless of civilian or military experience.

As our results demonstrate, operators can accept automation but still not trust it to high degrees. The overall average rating of trust could be perceived as a negative assessment of MAPP and indicative of an undesired level of distrust in our system. We propose that such a relationship is appropriate and desired in that operators find the automation useful, and even more so as they can express their risk levels, but are wary *enough* as not to over trust automation. Because heuristic algorithms are generally, but not always correct in navigation tasks, operators should never completely trust the algorithm. Automation bias, a human decision bias towards believing an automation-generated solution and not seeking disconfirming evidence, is a known negative consequence of too much trust in a system (Skitka, Mosier, & Burdick, 1999), which has been shown to be a problem in command and control systems (Cummings, 2004). Therefore, operators who completely trust a heuristic algorithm have an incorrect understanding of the automation's capabilities. This over-trust can lead to serious problems in these settings.

Designers of interfaces that rely on algorithms with degrees of solution uncertainty should consider both the consequences of too much and too little trust, and include features that reduce opacity, as well as communicate uncertainty in the automation's solutions. Moreover, when assessing trust through subjective surveys, designers should not necessarily desire the highest rating of complete trust in a system, especially when the underlying automation that drives the system is inherently flawed. Thus, we propose that designers of intelligent systems should seek some median level of acceptable trust, and that high and low ratings are both equally problematic.

## 5.0 Conclusion

Despite recent technological advancements in maritime navigation settings such as GPS-based navigation tools and electronic charts, the risk of maritime collisions and groundings is increasing. Navigation is a cognitively challenging task due to the multivariate nature of the problem and high uncertainty in the environment, which is especially true when navigating in coastal areas in time-pressured settings. We propose that the inclusion of automated path planners in these systems can significantly reduce workload and increase overall efficiency in time-pressured settings. However, because all algorithms carry their own degrees of uncertainty in addition to that present in the environment, such automated decision support tools must be designed to imbue appropriate levels of trust.



Based on an extensive cognitive task analysis, a Maritime Automated Path Planning tool was designed to primarily aid navigators in coastal and high traffic density time-pressured settings. This design included manual and automatic modes, with both enabling point and click interaction to easily create and modify a path free of obstacles such as shallow water or bad weather. An A* automatic path planner further reduced workload by providing a fully automated method for path generation that not only avoided static obstacles, but also dynamic contacts. Given that such planners may not always be correct due to their inherent brittleness, it is important that automated planners in such high uncertainty settings be designed specifically for transparency and explicit representation of uncertainty. As a result, MAPP was designed to engender user trust by allowing users to set constraint thresholds, and included visual representations of embedded uncertainties, such as the uncertainty associated with predicting other vessels' positions.

User testing showed that MAPP significantly improved the efficiency of the planning process, and promoted acceptance and appropriate trust. Despite early interviews that expressed distrust in automated decision support tools, military personnel statistically had no less trust than their civilian counterparts. More importantly, users accepted the automation and developed appropriate levels of trust in the heuristic-based automated path planner in that they gave the automation high marks for performance, but did not overwhelmingly trust the tool.

These results illustrate the importance of designing an automated decision support from principled information requirements, which reflect the cognitive needs of stakeholders. Moreover, while design interventions can be included to promote trust in a system, which is critical for the successful implementation of any automated system, it is important that designers understand the balance between too much and not enough trust in a system, especially when there are significant sources of exogenous uncertainty in high risk settings such as maritime navigation.



**Acknowledgements**

This work was sponsored by Rite Solutions Inc., Assett Inc., Mikel Inc., and the Office of Naval Research. We would also like to thank Northeast Maritime Institute, the MIT NROTC detachment, the crew of the *USS New Hampshire,* and the anonymous reviewers whose comments significantly improved the paper.

## References


Billings, C. E. (1997). *Aviation Automation: The Search for a Human-Centered Approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Carrigan, G. P. (2009). *The Design of an Intelligent Decision Support Tool for Submarine Commander.* MIT, Cambridge, MA.

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

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal Human-Computer Studies 58*, 697-718.

Endsley, M. R. (1993). *Situation Awareness and Workload: Flip Sides of the Same Coin.* Paper presented at the 7th International Symposium on Aviation Psychology, Columbus, Ohio.

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

Goldstine, H. H., Murray, F. J., & von Neumann, J. (1959). The Jacobi method for real symmetric matrices. *Journal of the Association for Computing Machinery*, 59-96.

Grabowski, M., & Sanborn, S. D. (2003). Human performance and embedded intelligent technology in safety-critical systems *International Journal of Human-Computer Studies, 58*(6), 637-670.

Guerlain, S. A., Smith, P. J., Obradovich, J. H., Rudmann, S., Smith, J. W., Svirbely, J., et al. (1999). Interactive Critiquing as a Form of Decision Support: An Empirical Evaluation. *Human Factors, 41*(1), 72-89.

Hart, P., Nilsson, N., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions of Systems Science and Cybernetics, 4*(2), 100-107.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Kelly, C., Boardman, M., Goillau, P., & Jeannot, E. (2003). *Guidelines for Trust in Future ATM Systems: Measures* (No. HRS/HSP-005-GUI-02). Brussels: EUROCONTROL

LaValle, S. M. (2006). *Planning Algorithms*: Cambridge University Press.

Layton, C., Smith, P. J., & McCoy, E. (1994). Design of a cooperative problem-solving system for en-route flight planning: An empirical evaluation. *Human Factors, 36*, 94-119.

Lee, J. D., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics, 35*(10), 1243-1270.

Lee, J. D., & Sanquist, T. F. (1996). Maritime automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance* (pp. 365-384). Mahwah, NJ: Lawrence Erlbaum Associates.

Lee, J. D., & Sanquist, T. F. (2000). Augmenting the Operator Function Model with Cognitive Operations: Assessing the Cognitive Demands of Technological Innovation in Ship Navigation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A, 30*(3), 273-285.

Lee, J. D., & See, K. A. (2004). Trust in technology: Designing for Appropriate Reliance. *Human Factors, 46*(1), 50-80.

MacMillan, J., Entin, E. B., & Serfaty, D. (1994). *Operator reliance on automated support for target acquisition.* Paper presented at the Human Factors and Ergonomics Society 38th Annual Meeting, Nashville, TN.

Montgomery, D. A., & Sorkin, R. D. (1996). Observer sensitivity to element reliability in a multielement visual display. *Human Factors, 38*, 484-494.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors, 39*(2), 230-253.

Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. Princeton, N.J: Princeton University Press.

Richardsen, P. W. (2008). *Expensive safety hangover in the shipping industry*. Oslo: Det Norske Veritas.

Rothgeb, M. (2008). *Course and Contact Advisory Agent to Manage Contacts*. Paper presented at the Undersea Human Systems Integration (HSI) Symposium.

Ruff, H., Narayanan, S., & Draper, M. H. (2002). Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles. *Presence, 11*(4), 335-351.

Russell, S., & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall.

Sauer, J., Wastell, D. G., Hockey, G. R. J., Crawshaw, C. M., Ishak, M., & Downing, J. C. (2002). Effects of display design on performance in a simulated ship navigation environment. *Ergonomics, 45*(5), 329-347.

Schraagen, J. M., Chipman, S., & Shalin, V. E. (2000). *Cognitive Task Analysis*. Mahwah, NJ: Erlbaum.

Schuffel, J. J., Boer, P. A., & van Breda, L. (1988). The ship's wheelhouse of the nineties: The navigation performance and mental workload of the officer of the watch. *Journal of Navigation, 42*(1), 60-72.

Sheridan, T. B. (1988). *Trustworthiness of command and control systems.* Paper presented at the *IFAC Man-Machine Systems*, Oulu, Finland.

Skitka, L. J., Mosier, K., & Burdick, M. D. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies, 51*, 991 - 1006.

Smierzchalski, R., & Michalewicz, Z. (1998). *Adaptive Modeling of a Ship Trajectory in Collision Situations at Sea.* Paper presented at the 2nd IEEE World Congress on Computational Intelligence, Alaska.

Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic Robotics*. Cambridge, MA: The MIT Press.

Uggirala, A., Gramopadhye, A. K., Melloy, B. J., & Toler, J. E. (2004). Measurement of trust in complex and dynamic systems using a quantitative approach. *International Journal of Industrial Ergonomics, 34*(3), 175-186

Urbanski, J., Morgas, W., & Czaplewski, K. (2008). The Ship's Navigation System for the XXI Century. *Transport Problems: an International Scientific Journal, 3*, 45-49.

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for Mental Workload Measurement in a Test and Evaluation Environment. *Human Factors, 35*(2), 263-281.

Winston., P. H. (1992). *Artificial Intelligence*. 3rd. ed. Boston: Addison-Wesley.

Xiao, J., Michalewicz, Z., Zhang, L.,, & Trojanowski, K. (1997). Adaptive Evolutionary Planner/Navigator for Mobile Robots. *IEEE Transactions on Evolutionary Computation, 1*, 18-28.