Single Operator, Multiple Robots: An Eye Movement Based Theoretic Model of Operator Situation Awareness

Raj M. Ratwani
Perceptronics Solutions Inc.
Arlington, VA
rajr@percsolutions.com

J. Malcolm McCurry, J. Gregory Trafton
Naval Research Laboratory
Washington, D.C.
{mccurry} {trafton} @itd.nrl.navy.mil

Abstract— **For a single operator to effectively control multiple robots, operator situation awareness is a critical component of the human-robot system. There are three levels of situation awareness: perception, comprehension, and projection into the future [1]. We focus on the perception level to develop a theoretic model of the perceptual-cognitive processes underlying situation awareness. Eye movement measures were developed as indicators of cognitive processing and these measures were used to account for operator situation awareness on a supervisory control task. The eye movement based model emphasizes the importance of visual scanning and attention allocation as the cognitive processes that lead to operator situation awareness and the model lays the groundwork for real-time prediction of operator situation awareness.**

Keywords- supervisory control; situation awareness; eye tracking; human-robot system

I. Introduction

As robots become cheaper and more capable, there is a strong desire to allow one human supervisor to control multiple robots. Several variables impact the effectiveness of the overall human-robot system, including the level of robot autonomy, the interface to the robots [2], and operator situation awareness.

The level of robot autonomy and the interface to the robots can impact how much monitoring the robot needs, called neglect time, and how much interface time it takes to interact with the robot, called interaction time [2]. The better the autonomy, the longer the robot can be neglected, so the longer the neglect time. When a robot needs attention by the human operator, the amount of time it takes to interact with that robot is the interaction time; a shorter interaction time is better than a long interaction time. A model was developed that specifies neglect tolerance to determine how robot autonomy and interface design interact to support supervisory control.

Reference [3] extended this work by putting a stronger emphasis on the person into their model. Specifically, they showed that when a vehicle needed attention and the operator ignores it, overall task performance was poorer than when they did notice it. They attribute this decrease in performance to the operator's lack of situation awareness [3].

Situation awareness (SA) is the perception of environmental elements within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future [1]. More specifically, there are three levels of SA that build upon each other. These three levels include perception, comprehension, and projection into the future. An operator may first notice that a vehicle has stopped moving (*perception*), understand that it needs some attention (*comprehend*), and predict that if nothing is done it will run out of fuel or be hit by another vehicle (*prediction*). These different levels of SA can be very valuable because breakdowns in each level may lead to different classes of solutions. For example, a breakdown in perception may be best handled by alerts and guidance, while a breakdown in comprehension may be facilitated by integration of multiple pieces of information, and breakdowns in prediction could be helped by predictive displays [4].

Situation awareness is typically measured by subjective measures [5], query methods [6, 7], and implicit performance measures [8]. While these methods are excellent for understanding situation awareness, facilitating cognitive engineering design, and improving training, they can be difficult to use for online measures of SA. Our goal is to not only understand the cognitive processes that lead to situation awareness, but also to be able to develop online measures of SA that can be used to predict when an operator is losing SA and provide facilitation to maintaining SA. One issue is to determine what level of SA would be the most productive to track and measure.

We believe that the first stage of SA, perception, is the most likely to be used to develop online measures of SA. There are several reasons why perception is probably the most productive of the three stages. First, the perception stage is clearly the first to occur; without perception, neither comprehension nor projection can take place. Second, the perception stage of SA has also been empirically shown to account for over 75% of pilot errors [9]. Hence, our focus is on the perception aspect of situation awareness.

A. Components of SA on a Supervisory Control Task

We focus on two different components that are important for maintaining situation awareness in a supervisory task. These two components are attention allocation and visual scanning.

1) Attention Allocation

Appropriately allocating attention to relevant cues in the environment is a critical component of SA. Without actually attending to relevant information cues the comprehension and prediction components of SA are not possible. Appropriate attention allocation entails attending to operationally relevant information [10-12] and avoiding distraction from information

This work was sponsored by work order number 55-8122 from the Office of Naval Research to Greg Trafton.

that is not currently operationally relevant. Attending to irrelevant information is often due to distraction because of a completely different task.

Highlighting the importance of proper attention allocation, [9] found that most errors associated with situational awareness occurred when all the necessary information was present, but when the operator failed to attend to the necessary information. Interestingly, situation awareness can be lost not only when working on a completely different task, but also when multi-tasking [13-15].

2) Visual Scanning

In a complex dynamic task, like supervisory control, operators have to determine what information is currently operationally relevant; visual scanning is one of the processes used to accomplish this goal. Visual scanning occurs when a person visually inspects an area for an object or event of interest. Scanning can be quite cursory (e.g., checking a rear-view mirror) or more involved (e.g., a radiologist examining an X-ray). Critically, visual scanning leads to encoding relevant objects and events, deciding whether to deal with an issue immediately, to come back to it some time in the future, or to ignore it. Deciding to come back to an event sometime in the future is a prospective memory event [16, 17]. These prospective memory events are interleaved with other task-relevant actions and have a strong mix of perception (scanning) and memory. Interestingly, memory for visual scanning events decays with time [18]. In a supervisory control task, an operator must continually scan for new issues (e.g., a vehicle that needs a new destination), deal with high priority events (e.g., engaging a vehicle once it reaches its destination), and check on previous events that were delayed (e.g., seeing if a vehicle is still on a collision course and may need help sooner).

Other researchers have suggested that visual scanning is a key component to SA. The SEEV model, for example, suggests that visual scanning is made up of saliency, is inhibited by the effort to move attention, and the expectation of seeing something valuable [19]. Having good visual scanning habits is correlated with noticing anomalous events [20] and better driving [21].

B. Developing Objective Measures of Situation Awareness

Our approach to developing objective measures of the cognitive processes underlying situation awareness was to focus on the pattern of operator's eye movements as a simulated supervisory control task was performed. A performance based approach [22] was used in this study whereby the outcome of particular discrete events on the supervisory control simulation served as a local assessment of good or poor operator situation awareness. Combining these local assessments comprises a global measure of situation awareness. The eye movements associated with these discrete events were analyzed; several different eye movement measures have been shown to be indicators of cognitive processing [23-25]. We developed eye movement measures that focused on operator attention allocation and visual scanning. Relying on eye movement measures as indices of cognitive process allows for the eventual development of real-time prediction of operator situation awareness.

There are several other methodologies that have been used to index the cognitive processing related to situation awareness; for example, psycho-physiological measures [26, 27] have been used to determine the cognitive components of situation awareness. Compared to the other methods, the eye movement methodology is nonintrusive and the data allow for measuring cognitive process online.

II. EXPERIMENT

To examine the cognitive processes underlying operator situation awareness in a supervisory control task, data were collected from a complex dynamic supervisory control simulation. In the simulation, a participant acted as a single operator controlling five semi-autonomous, homogenous uninhabited air vehicles (UAVs). The high-level goal of the simulation was to direct UAVs to specific targets on a map and visually identify key items at the target site. While participants performed the simulation eye movement data were collected.

A critical component to successfully completing the simulation was to prevent UAVs from hitting hazard areas, which dynamically moved around the map. If a UAV hits a hazard area the UAV takes damage and can become incapacitated. For an operator to avoid hazards, the operator must maintain situation awareness. First, operators must keep track of the positions of UAVs, targets, and hazards (perception). Second, the operator must retain these positions and comprehend the consequences of their locations (comprehension). Finally, the operator must be able to anticipate future consequences (projection); for example, anticipating that a UAV moving along a particular trajectory will eventually hit a hazard in the path and plan accordingly.

Each time a UAV's path intersected a hazard area the operator had to make an explicit action to divert the UAV and prevent damage. Whether or not the operator acted to prevent damage was used as a performance based measure of situation awareness. If the operator failed to make an action and the UAV hit the hazard area and took damage this was an indicator of poor situation awareness. If the operator made an action to prevent the UAV from hitting the hazard area this was an indicator of good situation awareness.

Each of the path-intersect-hazard events was treated as a discrete event and the eye movements associated with the event were analyzed. To develop a theoretic model of situation awareness logistic regression was used. A simple description of logistic regression is that it is a multiple linear regression model with a dichotomous variable as an outcome variable; a more detailed description can be found in [28]. The outcome of the path-intersect-hazard event (i.e. UAV does not take damage or UAV takes damage) was the dichotomous dependent variable.

Three predictors of situation awareness were developed that were centered on capturing attention allocation and visual scanning. Given the empirical data suggesting that the majority of situation awareness errors were associated with distraction and attention being allocated to other tasks [9], one predictor indicated whether the operator engaged a secondary task (i.e. a completely different task) during the UAV intersection event. Engaging in a secondary task during the UAV intersection event suggests that the operator will be less likely to have good situation awareness and there should be a greater likelihood of the UAV hitting the hazard since attention is being allocated elsewhere.

The other two predictors were based on operator's eye movements. The first eye movement measure was developed to capture attention being allocated to objects and events that were not directly relevant to the specific hazard event (e.g., multi-tasking). This predictor quantified the amount of time and attention spent on other subtasks of the task. Specifically, the off task measure was the number of fixations to off event cues; a greater number of off task fixations would suggest less operator situation awareness and a greater likelihood of the UAV taking damage.

The second eye movement measure was developed to capture the process of visual scanning on relevant information cues. This predictor was the frequency of fixations to event relevant cues. In addition to noticing relevant cues, the operator must keep relevant information in memory by scanning frequently. The nature of the simulation is such that there are several events happening in parallel and the operator must prioritize and may handle a more pressing event first, with the intention of returning to the UAV intersection event at a later time. Keeping a visual object in memory has been associated with re-fixations on the particular object [18]. Thus, to maintain the relevant cue in memory the operator should frequently scan and fixate on the cue while handling other events in the simulation. The frequency of relevant cue fixations predictor captures attention allocation to the relevant cues as well as the process of maintaining the information in memory; frequent scanning to a relevant cue will keep the event in memory, while less frequent scanning suggests that the event has been forgotten, leading to lower SA and a high likelihood of the UAV hitting the hazard.

A. Method

1) Partcipants

Thirteen George Mason University undergraduate students participated for course credit.

2) Simulation Description

The supervisory control task, originally designed as the Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles (RESCHU) [29, 30], was modified to only include homogenous UAVs. The interface of the supervisory control simulation has three main sections: the map window, the status window, and the payload window. The map area (Fig. 1, right side) displays UAVs (the numbered half ovals), targets (red diamonds) which UAVs should be directed to, and hazards (yellow circles) which should be avoided. The status window (Fig. 1, lower left corner) shows the status of the UAVs and includes information on vehicle damage, time until the vehicle reaches a waypoint or target, and time remaining in the simulation. The payload window (Fig.1, upper left corner) is used for a visual acquisition task (described below) which is performed when a UAV reaches a target and the target is engaged by the operator.

The operator's high level goal in the simulation is to direct UAVs to specific target areas, engage the targets, and perform a visual acquisition task once the UAV has engaged the target. The visual acquisition task is a different task from the map task and requires the participant to search for a predefined object in the payload window and to identify the object; on average the task took six seconds to complete. During the visual acquisition payload task the operator cannot make any actions on the map window of the interface.

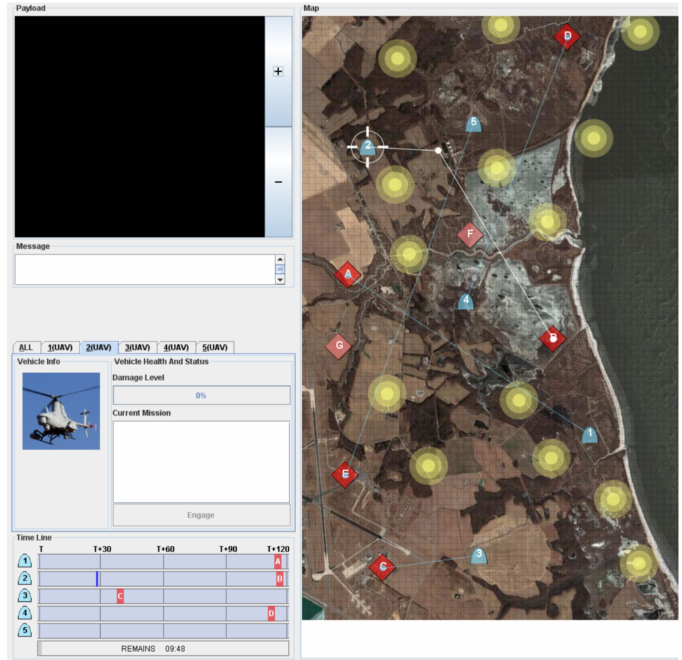

Figure 1. Screenshot of the supervisory control simulation.

During the simulation, hazard areas, which damage UAVs if passed through, dynamically appear on the map. If a UAV passes through a hazard area the UAV slows down and eventually the UAV will be incapacitated. At any given time there are 14 hazard areas, one randomly selected hazard area moves every 5 seconds. To avoid a hazard area the operator can perform two possible actions. First, the operator can assign the UAV to a different target, changing the path of the UAV and avoiding the hazard area. Second, the operator can assign specific waypoints to the UAV to divert the UAV from the path of the hazard, but keeping the final destination intact.

There are always 7 targets present on the map. At the start of the simulation the UAVs are randomly assigned to targets; thus, the UAVs may not be directed towards the most optimal target. In addition, after a target is engaged and the visual acquisition task is complete, the UAV is randomly assigned to a new target which again may not be optimal. The simulation is a complex task with multiple events happening in parallel. More than one UAV can be waiting at a target for engagement and more than one UAV can be on a path where a hazard area lies ahead.

When performing the simulation, participants were scored on their performance. Participants scored points by correctly completing the visual acquisition tasks; thus, participants sought to engage as many targets as possible and correctly complete as many visual acquisition targets as possible. Participants could maximize their score in several ways. First, participants should select the most optimal routes for UAVs to get to targets. Second, participants should re-plan when needed to make sure a UAV is headed towards the closest target. Third, participants should seek to avoid hazard areas because enough damage to a UAV will render the UAV incapacitated. Avoiding hazard areas is particularly important for achieving a high score since a single incapacitated UAV can drastically reduce the number of targets that can be engaged in the simulation.

3) Design and Procedure

The appearance of targets and hazards on the simulation map were randomized with the constraint that targets and hazards could be no closer than 2° of visual angle from each other. Because UAVs arrived at targets and could possibly pass through hazard areas, the same criteria of separation could not be applied to UAVs.

Prior to beginning the experiment, participants completed an interactive tutorial that explained all aspects of the simulation. During the tutorial, participants learned the objective of the simulation, how to control the UAVs (assigning targets, changing targets, assigning waypoints), and how to engage a target and complete the visual acquisition task. Participants were also warned of the dangers of hazards and were instructed on how to avoid hazards. The tutorial lasted approximately ten minutes. After competing the tutorial participants were given 10 minutes to practice the simulation.

After completing the practice session, participants were seated 47cm from the computer monitor and were calibrated with the eye tracker. Participants then began the actual simulation which lasted for 10 minutes. Participants were instructed to engage as many targets as possible and to complete as many visual acquisition tasks as possible to maximize their score.

4) Measures

Keystroke and mouse data were collected for each participant. The data from the supervisory control task were segmented in to discrete events based on when a UAV's trajectory intersected a hazard (called a path-intersect-hazard event). Each event started the moment the UAV's trajectory intersected a displayed hazard and ended when the UAV actually hit the hazard and took damage or when an explicit action was made by the operator to prevent the UAV from hitting the hazard. For each path-intersect-hazard event, the distance between the UAV and the hazard area varied and, consequently, the amount of time until the UAV would hit the hazard varied. At times, the UAV was close to the hazard and it would be beneficial for the operator to divert the UAV immediately, while at other times the UAV was quite far from the hazard and it would be beneficial for the operator to handle more pressing issues first and then return to divert the UAV. For example, if there are 40 seconds until a particular UAV hits a hazard, the operator may want to address other path-intersect-hazard events or engage a target and then return to divert the UAV at a later time.

Eye track data were collected using a Tobii 1750 eye tracker operating at 60hz. A fixation was defined as a minimum of three eye samples within 30 pixels (approximately 2º of visual angle) of each other, calculated in Euclidian distance. The average fixation duration on the task was 210 ms. Three areas of interest were defined that were directly related to the path-intersect-hazard events. These areas of interest were the UAV itself, the target area in which the UAV was headed for, and the hazard which blocked the UAV's path to the target area. For each path-intersect-hazard event, fixations on the map that did not land on one of these areas of interest and fixations on the status section of the interface were categorized as off task fixations. Note that an off task fixation for one event may not have been off task for another concurrent event. Fixations to the payload task were not analyzed.

Each of the predictors in the logistic regression model was defined as follows:

- The payload engaged predictor is a dichotomous variable indicating whether the participant performed the visual acquisition task in the payload window during the intersection event (0 = no payload task active, 1 = payload task active).
- The off task fixations predictor is a count of the total number of fixations to areas of the map that were not explicitly related to the intersection event. Objects in the map that were explicitly related to the intersection event were the UAV, the hazard, and the target that were involved in the intersection event. Additionally, any fixations that occurred while moving a UAV to a different location (e.g., fixating to an alternate target area) were also classified as relevant.
- The UAV scanning frequency represents the average amount of time between fixations on the relevant UAV involved in the intersection event (continuous variable in seconds).

B. Results and Discussion

Across the thirteen participants there were 177 path-intersect-hazard events. The average event time was 16.6 seconds with the shortest event lasting 0.8 seconds and longest event lasting 76 seconds. Approximately 82% (145 events) ended with the operator making an explicit action to prevent the UAV from hitting the hazard area. Redirecting the UAV to avoid a hazard area was accomplished in one of two ways: by either assigning a waypoint to divert the UAV out of the path of the hazard or by changing the target that the UAV was headed for to avoid the hazard. Approximately 18% (32 events) ended with the UAV hitting the hazard and taking damage. There were three possible reasons for a damage event. First, the operator may have failed to realize that the UAV was on a path that intersected a hazard. Second, the operator may have realized the UAV was going to hit the hazard too late and could not successfully execute the necessary actions to avoid the hazard. Finally, the operator may have noticed the UAV was on a path to hit a hazard, but the hazard was far enough away that more pressing events took precedent. The operator may have intended to return to redirect the UAV, but failed to remember to return and the UAV hit the hazard.

1) Developing a Logistic Regression Model

In order to create a logistic regression model of the path-intersect-hazard events, the outcome of damage and no damage was coded as the binary dependent variable. Each of the three predictors of interest (payload engaged, off task fixations, and UAV scanning frequency) was formulated for each of the 177 intersection events; a logistic regression model was calculated from these data. Equation (1) shows the logistic regression equation from the analysis:

$$\text{Predicted Logit of Damage} = -3.1 + (1.8 \times \text{payload engaged}) + (0.02 \times \text{off task fixations}) + (0.06 \times \text{UAV scanning frequency}) \tag{1}$$

The overall the logistic regression equation was significant, $\chi2\ (2) = 41.94$, $p<.001$. The log odds of damage occurring to a UAV were positively related to the operator working on the payload task ($p<.001$). Similarly, the log odds of damage were positively related to the number of off task fixations ($p<.05$) and the frequency of UAV scanning ($p<.05$). The results of the logistic regression model are summarized in Table 1.

Fig. 2 is a graph of the logistic regression model that illustrates how the predicted probability of a damage event occurring changes based on the three predictors. The payload engaged predictor is illustrated by the two sets of solid and dashed lines. The solid lines represent hazard-intersect events where the operator performed the visual acquisition task in the payload during the event and the dashed lines represent hazard-intersect events where no visual acquisition task was performed in the payload window. The number of off task fixations is represented on the x-axis. Finally, three levels of the UAV scanning frequency predictor (1, 7, and 15 seconds) are illustrated on the graph. The y-axis shows the predicted probability of a damage event occurring. Finally, the dots on the bottom and top of the graph are the actual no damage and damage events, respectively, from all the participants.

TABLE I. LOGISTIC REGRESSION TABLE

Predictor	**β**	**SE β**	**Walds χ^2**	**Significance value (*p*)**
Constant	-3.1	0.04	-7.4	0.001
Payload Engaged	1.8	0.05	3.7	0.001
Off Task Fixations	0.02	0.007	2.3	0.05
UAV Scanning Frequency	0.06	0.02	2.3	0.05

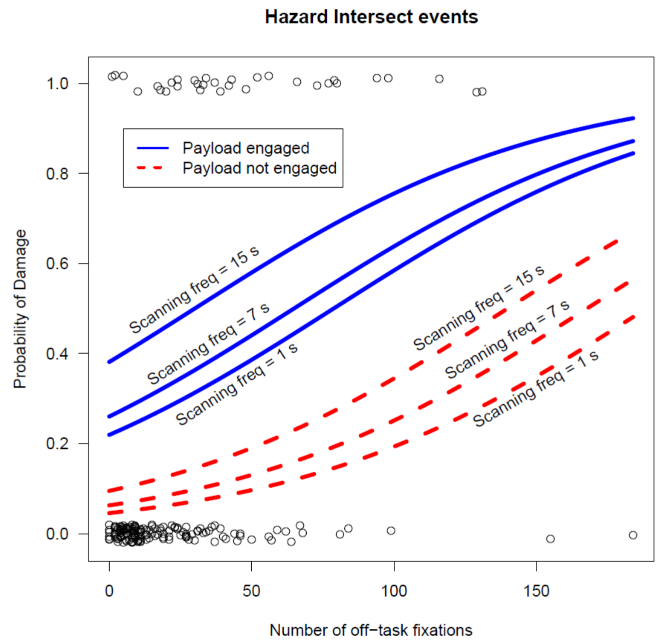

Figure 2. Logistic regression graph of the hazard-intersect events.

Several things are evident from Fig. 2. First, if the operator engages in the payload task the probability of the UAV taking damage is considerably higher than if the operator avoids engaging in the payload task. Second, the more the operator makes off task fixations the higher the probability that the UAV will take damage. Finally, infrequently scanning to the UAV that is involved in the intersection event suggests a higher likelihood of damage.

How well does the model fit the current data? The *c* statistic is a measure of model fit. This statistic represents the proportion of randomly selected hazard-intersection event pairs with different observed outcomes (i.e. no damage/damage) for which the model correctly predicts a higher probability for observations with the event outcome (i.e. damage) than the probability for nonevent observations (i.e. no damage). The *c* value for this logistic regression is .85, which means that for 85% of all possible pairs of intersection event actions, the model correctly assigned a higher probability to intersection events that resulted in damage than to intersection events that resulted in no damage. A *c* value of .85 is considered very good for logistic regression models.

The fact that each of the predictors loaded significantly in the logistic regression model and the strong model fit suggest that the cognitive processes that are represented by the predictors are accounting for operator situation awareness. The model suggests that attention allocation and visual scanning are key cognitive components to operator situation awareness.

2) Reciever-Operating Characteristic Analysis

While the *c* statistic provides a general measure of model fit, it is not clear how well the predicted probabilities of damage events generated from the logistic regression model match the actual data from the simulation. How many true damage events from the simulation dataset are actually predicted by the model?

To examine how well the model predicts the occurrence of actual damage and no damage events a receiver-operating characteristic (ROC) analysis was conducted. For each of the 177 hazard-intersect-events, the data associated with that event was entered in the logistic regression model. The model produced a predicted probability of a damage event occurring for each of the hazard-intersect events. These predicted probabilities were then compared to the actual occurrence of damage events to determine how accurate the model is.

Because the logistic regression model results in predicted probabilities, a threshold value must be established to categorize events as damage events or no damage events. Hazard-intersect events with probabilities that fall above the optimal threshold value will be categorized as damage events and hazard-intersect events with probabilities that fall below the threshold will be considered no damage events.

A ROC analysis provides a method for visualizing the performance of the logistic regression model at different threshold values [31]. In order to develop the ROC curves, these threshold values were systematically varied from 0 to 100 percent. The predicted damage and no damage events at each threshold value were compared to the actual data to generate the true positive and false positive rates. Each of these pairs of values was then used to generate the ROC curve seen in Fig 3.

The ROC curve in Fig. 3 is plotted in ROC space. Points that fall in the upper left hand corner represent perfect prediction; the points result in a high true positive rate and a low false positive rate. Thus, the threshold associated with the point closest to the upper left hand corner represents the optimal threshold for maximizing true positives and minimizing false positives. In this ROC analysis the threshold value associated with the point closest to the upper left hand corner is 14.7%.

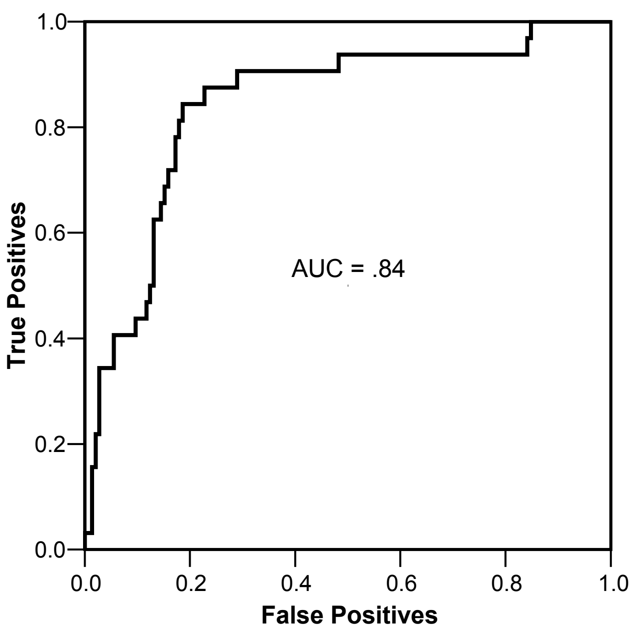

Figure 3. ROC curve for the logisitic regression model.

In order to quantitatively determine how robust the logistic regression model is at predicting damage and no damage events in this dataset, the area under the ROC curve can be examined. The area under the curve (AUC) represents the probability that the logistic regression model will rank a randomly chosen positive instance (i.e. a damage event) higher than a randomly chosen negative instance (i.e. no damage) [31, 32]. The AUC is similar to the *c* statistic that was used to determine model fit. The area under the ROC curve for the logistic regression model is .84. These values are considered very good for classification purposes and suggest that the logistic regression model is correctly ranking a large number of the damage events.

To examine the actual number of path-intersect-hazard events that the logistic regression model would correctly classify, the optimal threshold (14.7%) was used to categorize each of the predicted probabilities as damage or no damage events. The predicted outcomes were then compared to the true outcomes to determine the accuracy of the logistic regression model. The confusion matrix showing the performance of the model is displayed in Table 2.

TABLE II. CONFUSION MATRIX

	Actual Value	
Predicted Value	True Positive 84.4% (27)	False Positive 18.6% (27)
	False Negative 15.6% (5)	True Negative 81.4% (118)

The logistic regression model correctly classifies approximately 85% of the actual damage cases and correctly classifies approximately 81% of the no damage cases. Overall, the model performs well suggesting that the predictors, and the underlying cognitive processes represented by the predictors, are good indicators of situation awareness.

III. General Discussion

Previous theoreticians have emphasized the importance of attention allocation and visual scanning in situation awareness. We instantiated two different measures of attention allocation (off-task performance and multi-tasking), as well as a single measure of visual scanning, to account for the cognitive processes underlying situation awareness. All three SA measures were found to be significantly related to performance on a supervisory control task, supporting the theoretic importance of both attention allocation and visual scanning for maintaining situation awareness.

A. Visual Scanning

Previous researchers have suggested that visual scanning is important for maintaining SA [19-21]. However, previous empirical research did not directly measure visual scanning [21] or, when they did measure visual processing, it was as a percentage of looking at the relevant area [20]. By showing that visual scanning can be operationalized as the amount of time between relevant fixations, we have identified a strong measure of level 1 SA. Specifically, we found that if people frequently scan to an important visual event, they are likely to notice when that event needs operator attention.

There are several reasons why this measure is such a good measure of SA. First, in a dynamic task, different events need to be prioritized. In a supervisory control task, if a UAV is likely to need attention sometime in the future, but a different UAV needs attention now, the operator is likely to deal with the higher priority UAV first, but then will need to remember to take care of the less immediate event. Because memory is not perfect, we interpret periodic scans back to a UAV as a "memory refresh" [18].

Second, while periodic scans can serve a monitoring function, the operator must still remember to perform those scans or be extremely systematic in their scanning behavior. Unfortunately, people are typically driven more by saliency of objects than by systematic search processes [33].

Finally, even when the operator's task is relatively easy, periodic scanning is needed because at any point in time a UAV may need attention (e.g., it may be on a hazard intersect or it may have arrived at its goal). Again, frequent visual scanning allows operators to maintain high situation awareness and keep their UAVs on task with as little idle time as possible. Finally, frequent visual scanning can reduce wait time [3] and can allow more UAVs to be controlled [2].

B. Attention Allocation

We have also found additional support that two different aspects of attention allocation impacts situation awareness. In supervisory control tasks, performing secondary tasks, even if they are relevant to the mission, reduces situation awareness. Other researchers have shown that paying attention to other tasks can reduce SA as well [13-14].

We have also shown that multi-tasking can reduce situation awareness and increase the likelihood of a UAV taking damage. Other researchers have also shown that multi-tasking while driving can increase the rate of crashes [21].

It should be noted that both aspects of attention allocation are needed as part of any supervisory control task. When a UAV gets to its target location, the operator must tend to it – in our case, perform a payload task. This tending to, by definition, takes attention away from ongoing hazard-intersect events, reducing SA. Similarly, dealing with a higher priority event also leads to a reduction in SA for a hazard-intersect event that will occur further in the future. While it would be best, of course, to be able to immediately deal with a hazard event as soon as it happens, in a dynamic task, this simply is not possible. Thus, there exists an interesting tension between maintaining high situation awareness and actually accomplishing necessary aspects of the task. We discuss ways of dealing with this tension in a later section.

C. An integrated model of Situation Awareness

Critically, our approach has allowed us to explore these three measures of SA in combination. Most researchers focus on the impact of one of these measures by itself while holding the others constant experimentally. Our approach assumes that all of these components are important for situation awareness and they all occur in complex dynamic tasks. By examining them in an integrative model, we found that all of them show an impact on SA independent of each other. To the best of our knowledge, we are the first researchers to show precise, quantitative measures of both visual scanning and attention allocation in an integrated model.

D. Improving Operator Situation Awareness

Our theoretic model of situation awareness can be used to improve operator situation awareness in at least two ways. First, the model predicts the conditions under which operators are most vulnerable to a loss of situation awareness and the model specifies some of the critical cognitive components to maintaining situation awareness. This information can be integrated with current operator training procedures and this information can be leveraged to improve interface design.

Based on the model, supervisory control operators are most vulnerable to a loss of situation awareness when the operator is multitasking and when the operator decides to engage in off-task events. Training should stress the importance of frequent visual scanning, particularly when multitasking or working on an off-task event. Interfaces can be designed to facilitate operator scanning, especially when off-task events need to be worked on, by reducing the spatial distance between information displays.

A second way to leverage our theoretic model to improve operator situation awareness is to use the model to develop a real-time system that predicts the operators' level of situation awareness. Real-time feedback systems based on various online measures of cognitive process have been used in several domains [26, 27, 34, 35].

Using our approach, operator's eye movements can be analyzed in real-time as the operator is controlling multiple robots and the predicted probability of operator situation awareness failure can be continually calculated. Given that there are multiple robots multiple predictive models can be run simultaneously. If the predicted probability reaches a threshold that suggests the operator is losing situation awareness the operator can be immediately alerted.

A real-time system that can predict when an operator is losing SA can be particularly useful to combat problems associated with automation. For example, the operator-out-of-the-loop performance problem can be addressed head on by a real-time prediction system. The operator-out-of-the-loop problem [36, 37] occurs when operators rely on automation and reduce their monitoring of the system. If a problem arises, operators don't know the state of the system and have difficulty assessing and addressing the problem. With a real-time system that predicts when operators are losing SA, the operator can be alerted and can then attempt to maintain appropriate visual scanning to prevent the loss of SA. The theoretic model of operator situation awareness developed here lays the groundwork for the development of such a system.

Acknowledgment

The authors thank Missy Cummings and the Human Automation Lab at MIT for providing us with RESCHU. The authors also thank Richard Louie for providing assistance in conducting the experiment.

References

[1] M. R. Endsley, "Toward a theory of situation awareness in dyanmic systems," *Human Factors*, vol. 37, pp. 32-64, 1995.

[2] J. W. Crandall, M. A. Goodrich, D. R. Olsen, and C. W. Nielsen, "Validating human-robot interaction schemes in multitasking environments," *Systems, Man, and Cybernetics*, vol. 35, pp. 438-449, 2005.

[3] M. L. Cummings and P. J. Mitchell, "Predicting controller capacity in supervisory control of multiple UAVs," *IEEE Systems, Man, and Cybernetcis, Part A: Systems and Humans*, vol. 38, pp. 451-460, 2008.

[4] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "Situation awareness, mental workload, and trust in automation: Viable , empirically supported cogntive engineering constructs," *Journal of Cognitive Engineering and Decision Making*, vol. 2, pp. 140-160, 2008.

[5] R. M. Taylor, "Situational Awareness rating technique (SART): The development of a tool for aircrew systems design," in *AGARD Conference Proceedings*, 1990, pp. 3/1-3/17.

[6] M. R. Endsley, "Measurement of situation awareness in dynamic systems.," *Human Factors*, vol. 37, pp. 65-84, 1995.

[7] F. T. Durso and A. R. Dattel, "SPAM: The real-time assessment of SA," in *A cognitive approach to situation awareness: Theory and application*, S. Banbury and S. Tremblay, Eds., 2004, pp. 137-154.

[8] A. D. Andre, C. D. Wickens, and L. Moorman, "Display formatting techniques for improving situation awareness in the aircraft cockpit," *The International Journal of Aviation Psychology*, pp. 205-218, 1991.

[9] D. G. Jones and M. R. Endsley, "Sources of situation awareness errors in aviation.," *Aviation, Space and Environmental Medicine*, vol. 67, pp. 507-512, 1996.

[10] M. R. Endsley and R. P. Smith, "Attention distribution and decision making in tactical air combat," *Human Factors*, vol. 38, pp. 232-249, 1996.

[11] M. R. Endsley and M. D. Rodgers, "Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errorsand automation," *Air Traffic Control Quarterly*, vol. 6, pp. 21-44, 1998.

[12] L. Gugerty, "Evidence from a partial report task for forgetting in dynamic spatial memory," *Human Factors*, vol. 40, pp. 498-508, 1998.

[13] L. Gugerty, M. Rakauskas, and J. Brooks, "Effects of remote and in-person verbal interactions on verbalization rates and attention to dynamic spatial scenes," *Accident analysis and prevention*, vol. 36, pp. 1029-1043, 2004.

[14] L. Gugerty, "Situation awareness in driving," in *Handbook for Driving Simulation in Engineering Medicine and Psychology*, J. Lee, M. Rizzo, D. Fisher, and J. Caird, Eds.: CRC Press, in press.

[15] W. J. Horrey, C. D. Wickens, and K. P. Consalus, "Modeling drivers' visual attention allocation while interacting with in-vehicle technologies," *Journal of Experimental Psychology: Applied*, vol. 12, pp. 67-78, 2006.

[16] G. O. Einstein, M. A. McDaniel, and M. Brandimonte, *Prospective memory: Theory and application*: L. Erlbaum, 1996.

[17] J. G. Trafton, E. M. Altmann, D. P. Brock, and F. E. Mintz, "Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal," *International Journal of Human-Computer Studies*, vol. 58, pp. 583-603, 2003.

[18] G. J. Zelinsky and L. C. Loschky, "Using eye movements to study working memory rehearsal for objects in visual scenes," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, N. Taatgen, H. V. Rijn, and L. Schomaker, Eds., 2009.

[19] C. D. Wickens, J. Goh, J. Helleberg, and W. J. Horrey, "Attentional models of multitask pilot performance using advnaced display technology," *Human Factors*, vol. 45, pp. 360-380, 2003.

[20] L. C. Thomas and C. D. Wickens, "Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display," in *Proceedings of the 48th Annual Meeting of the Human Factors Society*, 2004.

[21] L. Gugerty, "Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory," *Journal of Experimental Psychology:Applied*, vol. 3, pp. 42-66, 1997.

[22] P. Salmon, N. A. Stanton, G. Walker, and D. Green, "Situation measurement: A review of applicability for C4i environments," *Applied Ergonomics*, vol. 37, pp. 225-238, 2006.

[23] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 85, pp. 618-660, 1998.

[24] K. Rayner and R. K. Morris, "Do eye-movements reflect higher order processes in reading?," in *From eye to mind. Information aquisition in perception, search, and reading*, R. Groner, D. d'Ydewalle, and R. Parham, Eds. Amsterdam: North-Holland, 1990, pp. 191-204.

[25] M. A. Just and P. A. Carpenter, "Eye fixations and cognitive processes," *Cognitive Psychology*, vol. 8, pp. 441-480, 1976.

[26] G. F. Wilson and C. A. Russell, "Performance enhancement in an unihabited air vehicle task using psychophysiologically determined adpative aiding," *Human Factors*, vol. 49, pp. 1005-1018, 2007.

[27] G. F. Wilson and C. A. Russell, "Operator functional state classification using multiple psychophysiological features in an air traffic control task," *Human Factors*, vol. 45, pp. 381-389, 2003.

[28] C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," *The Journal of Educational Research*, vol. 96, pp. 3-14, 2002.

[29] Y. Boussemart and M. L. Cummings, "Behavioral recognition and prediction of an operator supervising multiple heterogeneous unmanned vehicles," in *Humans operating unmanned systems*, 2008.

[30] C. Nehme, "Modeling human supervisory control in heterogeneous unmanned vehicle systems," in *Deptartment of Aeronautics and Stronautics*. Cambridge, MA: MIT, 2009.

[31] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861-874, 2006.

[32] N. A. Macmillan and C. D. Creelman, *Detection Theorys: A User's Guide*: Lawrence Erlbaum Associates, 2005.

[33] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, pp. 1489-1506, 2000.

[34] R. M. Ratwani, J. M. McCurry, and J. G. Trafton, "Predicting postcompletion errors using eye movements," in *Proceedings of the 2008 Conference on Human Factors in Computing Systems, CHI 2008*, M. Czerwinski, A. M. Lund, and D. S. Tan, Eds.: ACM, 2008, pp. 539-542.

[35] R. M. Ratwani and J. G. Trafton, "Developing a Predictive Model of Postcompletion Errors," in *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*, 2009.

[36] M. R. Endsley and E. O. Kiris, "The out-of-the-loop performance problem and level of control in automation," *Human Factors*, vol. 37, 1995.

[37] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and level sof human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and human*, vol. 30, pp. 286-297, 2000.