Situation Awareness Reacquisition in a Supervisory Control Task

Daniel Gartenberg[1,2], Malcolm McCurry[2], Greg Trafton[2]

dgartenb@gmu.edu
George Mason University[1]
Fairfax, VA

mccurry@itd.nrl.navy.mil, greg.trafton@nrl.navy.mil
Naval Research Laboratory[2]
Washington, DC

External interruptions, task switching, distractions, and multi-tasking can all affect situation awareness (SA). This study focused on how SA is reacquired after a brief task-related break. Participants controlled multiple unmanned aerial vehicles, avoiding hazards and navigating vehicles to their target destination. In the dual-task condition, participants completed a payload sub-task after engaging a vehicle for mission completion. In the single task condition, participants did not complete the additional sub-task after engaging a vehicle for mission completion. Patterns of eye fixations were found that characterized instances when SA was being reacquired (dual task) and instances when there was continuous task performance (single task). After a task break, SA was reacquired by quickly scanning a diverse group of objects that had been previously looked at. When there was no task break, participants slowly fixated on a few objects that were novel. We interpret these findings as suggesting that when SA needs to be reacquired, previous goals and plans need to be reinstated, while during normal task behavior, participants seek novel and changing events. These findings support the Memory for Goals (MFG) model and the integrated framework for maintaining and recovering SA. We discuss implications for developing process models that evaluate SA in real-time.

INTRODUCTION

An air-traffic controller is responsible for monitoring the trajectories of multiple aircrafts, which can at times be on path towards hazardous events. To respond to a hazardous event, the operator must redirect focus from monitoring the various aircrafts to preventing the hazardous event from occurring. The operator directs the pilots of the aircraft to change course, which can have the effect of preventing a crash. However, the operator has spent a great deal of time and cognitive resources on addressing the hazardous event, and this has come at a cost. The other vehicles on the screen have not been looked at for the entire time that the hazardous event was being addressed, thereby reducing the operator's situation awareness (SA) of those other vehicles. Pilots, automobile drivers, and power-plant workers have similar jobs as the air-traffic controller because their jobs require making fast paced decisions after addressing subtasks. We investigated the perceptual and cognitive processes that characterize periods of SA reacquisition after completing such subtask.

There are many methods of determining someone's level of SA. These methods range from freeze paradigms, like the situation awareness global assessment technique (SAGAT) (Endsley, 1995a), real time probes, like the situation present assessment method (SPAM) (Durso, Truitt, Hackworth, Crutchfield, Nikolic, et al., 1995), and process indices, such as eye movements (Salmon, Stanton, Walker, & Green, 2006).

There are clear strengths and weaknesses between these measures. A strength of SAGAT and SPAM is that they directly measure SA, yet a weakness is that they require responding to questions during the task. This can pose an additional burden on the operator upon resumption, since these tasks are dynamic. A strength of process indices, like eye movements, is that they are noninvasive and do not impact task performance. Yet a weakness is that it is difficult to quantify a relationship between eye movements and SA (Salmon, et al, 2006).

We used eye movements to characterize the situation after taking a brief break from a primary task. Task breaks can occur due to an external interruption, task switching, distraction, or even multi-tasking. While these breaks can occur for different reasons (e.g., internal decisions or the external environment), all of these breaks have been shown to decrease performance in terms of both increasing the time it takes to act, or the resumption lag (Hodgetts & Jones, 2006a; Monk, Boehm-Davis, & Trafton, 2004; Trafton, Altmann, Brock, & Mintz, 2003; Altmann, & Gray, 2008), and the number of errors that are made (Ratwani, McCurry, & Trafton, 2008). In a dynamic task, operators may need to reacquire SA after a break in order to understand what (if anything) has changed since they were otherwise engaged.

To understand SA reacquisition, we adapted the Memory for Goals (MFG) model. The MFG model was initially used to interpret how subgoals are suspended and resumed in a problem-solving task (Altmann & Trafton, 2002). It has since been used to interpret the time costs of interruptions (Altmann & Trafton, 2007) and the factors affecting sequence error (Trafton, Altmann, & Ratwani, 2009).

The MFG model is based on the hypothetical construct of activation of memory elements—in particular, activation as construed in the ACT-R (Adaptive Control of Thought-Rational) cognitive theory (Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004). The MFG model inherits two basic processing assumptions from ACT-R. The first is that when central cognition queries memory, memory returns the item that is most active at that instant. The second is that the activation of a given memory element fluctuates noisily from moment to moment about a mean value. The MFG model makes additional assumptions that govern the activation of fine-grained episodic memory codes; these codes can be considered task-related subgoals. The first assumption is that, after an episodic code is encoded, its activation automatically decays, such that a retrieval request is more likely to return a recent code than a less-recent code (all else being equal). The second assumption is that an episodic code can also be primed by contextual retrieval cues, thereby overcoming effects of decay (i.e. the priming constraint) (Altmann & Trafton, 2002).

The MFG model can be leveraged to make predictions about SA reacquisition. After a break from a task, people will need to examine the environment to determine what has changed, perhaps searching out current high priority tasks. MFG predicts that the time it takes to resume after a break will be longer than the situations where there is not a break (Trafton et al., 2003). An ancillary prediction is that there will also be more fixations when operators attempt to reacquire SA.

MFG also makes a more nuanced prediction based on the priming constraint about where those fixations should be. If operators need to reinstate previous context (i.e. reacquire SA), the fixations themselves should be on objects that they have looked at recently. The reason for this is that by looking at old objects, priming can flow from these objects to help reinstate previous goals and plans. In effect, these old objects are acting as contextual cues for SA reacquisition.

Distinct from the SA reacquisition stage is the situation when the goal state is highly activated. St John and Smallman (2008) developed an integrated framework for maintaining and recovering SA that uses concepts from the MFG model to make predictions about how attention is allocated in dynamic tasks when the goal state is degraded and when it is highly activated. St John and Smallman (2008) organized SA into four stages: 1) change detection before an interruption, 2) preparing for an interruption, 3) reorienting after an interruption, and 4) change detection that occurred after an interruption. Change detection during recovery (i.e. after an interruption) requires the additional step of reorientation. This is similar to the contextual cueing that occurs in the MFG model when the goal state is degraded. Yet the change detection stage before an interruption does not have the additional step of reacquisition of the goal state because it has not been degraded by the interruption. Therefore, attention is allocated to detecting novelty in the dynamically changing environment when there is no task interruption.

The integrated framework for maintaining and recovering SA and MFG suggest that situations when SA is high will be characterized by increased fixations to novel stimuli, and that when SA needs to be reacquired, more fixations will be on objects that were previously looked at. The goal of this paper was to determine if situations where SA is relatively lower, due to a task break, can be differentiated from situations where there was no such task break, and if these fixation patterns are consistent with the MFG model.

METHOD

Participants

Eighty-one George Mason University undergraduate students participated for course credit. All participation was voluntary. There were sixty females and twenty-one males who participated in the study. The average age of participants was 20.43 years old with a standard deviation of 2.96 years old. Participants were asked to rate how often they played video games on a scale of 1 (never), 2 (sometimes), or 3 (a lot). The average amount of video game play was 1.80 with a standard deviation of 0.69. All participants had normal or corrected-to-normal vision.

Eye data for three participants were eliminated because it was not accurately captured. In total, seventy-eight participants were analyzed, thirty-eight in the interruption condition and forty in the no interruption condition.

Materials

The supervisory control simulation was the Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles (RESCHU) (Boussemart and Cummings, 2008). The task involved navigating homogenous unmanned aerial vehicles (UAVs) in a dynamically changing environment.

The simulation consisted of three main sections: a map window, a payload window, and a status window (see Figure 1). The map area displayed UAVs (blue half ovals), targets (red diamonds), which UAVs should be directed to, and hazards (yellow circles), which should be avoided. The payload window (top left) displayed a visual search screen where the participant was instructed to identify an object as part of a payload operation (described later). The status window (bottom left) depicted a timeline of when the UAVs would reach important events, which included waypoints and the target of the UAV.

The task began with five UAVs moving at a fixed speed. The UAVs continued to move at this fixed speed throughout the duration of the task. Eighteen hazard areas moved randomly every four-seconds, with the constraint that the hazards could not appear closer than 3° of visual angle (about 50 pixels) away from any UAV. If the UAV passed through a hazard, damage occurred. Damage was indicated as a bar in the status window. The appearance of targets and hazards on the simulation map were randomized with the constraint that targets and hazards could be no closer than 3° of visual angle from each other. This assured that targets and hazards could not co-occur in the same space. There were always 7 targets present on the map.

The operator's goal was to direct UAVs to target areas, while avoiding hazard areas. To avoid a hazard area the operator could assign the UAV to a different target or the operator could assign specific waypoints to the UAV, which effectively allowed the operator to pilot the UAV around hazard areas. At the start of the simulation the UAVs were randomly assigned to targets. Once the UAV reached the target destination, the target flashed red until it was engaged. When the UAV arrived at the target the participant then could right click on the vehicle and engage it.

14:34:07 CORRECT!, SCORE!
14:34:24 [MISSION(2)] Find the blue area on the back of the boat

Find the blue area on the back of the boat

Figure 1. Supervisory control simulation

In the single-task condition, simply engaging the vehicle completed the mission. In the dual-task condition, engaging the vehicle triggered the payload, where the participant performed a visual search task to identify an object, such as a ship or a car.

During the payload, the vehicles in the map task continued to move toward their corresponding target. After identifying the object in the payload task, the UAV's mission was completed. Completion of the payload ranged between 1.3 – 56.8 seconds, with a mean of 4.8 seconds and standard deviation of 4 seconds. After completing the mission, the UAV was randomly assigned to a new target that did not have any other UAV assigned to it.

The simulation is a complex task with multiple events occurring in parallel. Each simulation had unique characteristics with randomly generated trajectories, locations, and objectives, similar to a real-world environment. More than one UAV could be waiting at a target for engagement, multiple UAVs could be on a path to a hazardous area, and it was left to the participant's discretion to act on any one of the five vehicles.

Design and Procedure

The experiment was a between groups design. One group was assigned to the single task condition and the other group was assigned to the dual task condition. The dependent variables were related to the patterns of eye fixations directly after participant completed a vehicle engagement and were either required to complete the payload task (dual task condition) or were not required to complete the payload task (single task condition.

Prior to beginning the experiment, both groups completed an interactive tutorial that explained all aspects of the simulation. Participants learned about the objective of the simulation, which was to prevent as much damage as possible and engage as many vehicles as possible. Additionally, participants learned how to control the UAVs (assigning targets, changing targets, assigning waypoints) and to engage a target (by right clicking on the target and pressing the engage pop-up window). Participants were also warned of the dangers of hazards and were instructed on how to avoid hazards. The tutorial lasted approximately ten minutes.

Then the participants began a practice session where they were exposed to the task for which they were assigned. During this practice session the participant was instructed to interact with the task until the experimenter felt that they understood the task and could complete each sub-activity.

After completing the practice session, participants were seated approximately 66cm from the computer monitor and were calibrated on the eye tracker. Participants then began the simulation, which lasted for 10 minutes. Participants were again instructed to maximize their score by engaging as many targets as possible and preventing as much damage as possible. When the simulation ended, participants received feedback on how many vehicles they engaged and total vehicle damage. Participants were re-calibrated and were run in a second 10-minute session. Participants were run in the same condition for both sessions and both sessions were combined in the analysis.

Measures

Keystroke and mouse data were collected for each participant. Eye tracking data were collected using an SMI eye tracker operating at 250hz. A fixation was defined as a minimum of fifteen eye samples within 50 pixels (approximately 3° of visual angle) of each other, calculated in Euclidian distance.

Segmenting the task into intervals of interest. One way to analyze a dynamic continuous task is by distinguishing between interaction time intervals - where participants make actions, and wait time intervals (i.e. monitoring intervals) - where participants monitor the screen and decide what to do next (Crandall, Goodrich, Olsen, & Nielsen, 2005). A monitoring interval of particular interest was the instance of time after the mission complete action because this interval preceded a break in the dual-task condition or no break in the single-task condition. Analyzing the intervals between actions was similar to the technique used in Altmann and Trafton (2004), who examined monitoring intervals that either preceded an interruption or did not precede an interruption. Instead of comparing the interval of time before an interruption, we compared the interval after either completing a payload subtask or not completing a subtask. On average the participant completed 54.3 missions across the two 10 minute sessions, resulting in an average of 54.3 monitoring intervals analyzed per participant.

Categorizing fixations. The pattern of fixations after a mission was completed and before the next action was analyzed. A mission was completed after a vehicle arrived at its target location and the participant finished engaging the vehicle (pressing the engage button in the single task condition or pressing the engage button and completing the payload in the dual task condition). There were a total of five UAVs on the screen, which had different targets and hazards associated with them. A fixation on a vehicle, the vehicle's relevant hazard, and / or the vehicle's relevant target was classified as a 'unique object fixation'. Since there were a total of five vehicles, there was a maximum of five unique object fixations in the after mission completion interval.

Unique object fixations were further characterized as either re-fixations or novel fixations. A re-fixation occurred if the fixation that happened after the mission completion action was on a situation that the participant had fixated on in the past, prior to the mission completion action. The time interval for which participants could fixate on objects in the past varied based on how long the situation (i.e. hazard intersect or vehicle waiting to be engaged event) occurred. In other words, if a fixation that occurred during the monitoring interval after the mission completion was on a situation that had been looked at in the past, then the fixation was characterized as a re-fixation. A novel fixation occurred when a fixation in the monitoring interval after the task completion was on an object situation that was never fixated on in the past. For example, assume that a participant looked at vehicles 1, 2, and 3 and then engages vehicle 4. After the vehicle 4 mission is completed, the participant looks again at vehicle 1 (a re-fixation) and at vehicle 5 (a novel fixation).

RESULTS AND DISCUSSION

All analyses were conducted on the monitoring interval of time after the mission completion and before the next action. Consistent with MFG, participants in the dual-task condition took longer to resume (M = 5033.17 ms, SD = 1144.58 ms) than participants in the single-task condition (M = 4124.23 ms, SD = 1121.55 ms), $t(76) = 3.54$, $p < .05$, $d = .80$. This is consistent with other research showing that there is an increased resumption lag after an interruption (Hodgetts & Jones, 2006a; Monk, et al., 2004; Trafton, et al., 2003.)

Another interesting difference between the single task and the dual task was that the fixation duration was shorter in the dual task condition (M = 313.70 ms, SD = 69.46 ms) than the single task condition (M = 403.56 ms, SD = 81.72 ms), $t(76) = 5.24$, $p < .05$, $d = 1.19$. This showed that perceptual processes play a role in distinguishing between the interval after the single task and the interval after the dual task reacquisition.

A mixed model ANOVA was run to examine the relationship between novelty/re-fixation and single/dual task. Task condition was the between groups factor and fixation type (re-fixation or novelty fixation) was the within groups factor. The dependent measure was the number of unique fixations that occurred in the monitoring interval after the mission completion and before the next action.

Consistent with MFG, there was a main effect of task condition, with more unique object fixations in the dual-task condition (M = 2.57, SD = 0.44) than the single task condition (M = 2.05, SD = 0.25), $F(1, 76) = 39.38$, $p < .05$, $\eta^2 = 0.34$. There was no main effect of fixation type $F(1, 76) = 0.44$, $p = .51$, $\eta^2 = 0.01$. Consistent with the MFG model and the integrated framework for maintaining and recovering SA, there was an interaction between task condition and fixation type, $F(1, 76) = 11.58$, $p < .05$, $\eta^2 = 0.13$. As suggested in Fig 2, by the unequal variances between conditions, we used the Welsh's correction to account for these differences in variance. All groups were significantly different from all other groups ($p < .05$ using the Bengamini Hachberg correction). As Fig 2 suggests, in the dual task condition there were more re-fixations than novelty fixations, but in the single-task condition there were more novelty fixations than re-fixations.

This nuanced finding was based on the Memory for Goals model. MFG suggests that after a break, the goal representation needs to be reinstated, which can occur through either internal (memory, imagination) or external (fixation) cues. Reinstating a representation primes the participant to remember those previous goals. We interpret this finding as participants using contextual cues to increase activation of the goal state during SA reacquisition. Yet during continuous task execution, participants sought out novel changes in the task.

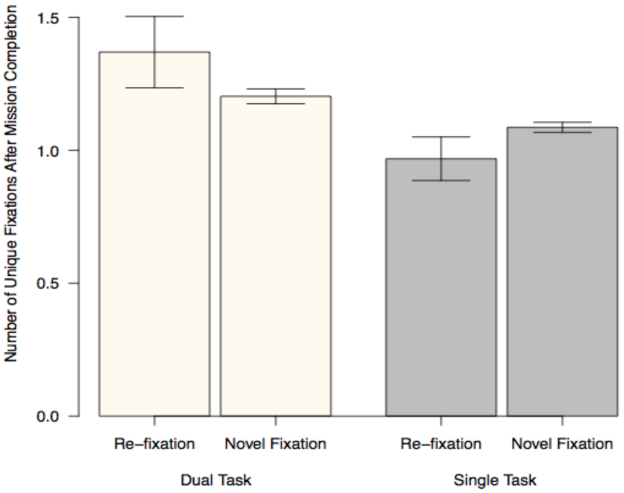

Figure 2. *Type of Fixations Based on Task Condition. Error bars are 95% Confidence intervals*

CONCLUSION

This study focused on how people reacquire SA after a brief task-related break. Participants were run through a dynamic supervisory control task. In one condition (the dual-task), participants were required to complete a sub-task after engaging a vehicle for mission completion. In the other condition (single task), participants did not complete the additional sub-task after engaging a vehicle for mission completion. SA reacquisition was characterized by more object fixations, faster fixations, and more re-fixations to known objects than novel objects. In contrast, when SA reacquisition was not needed, there were fewer object fixations, faster fixations, and more fixations on novel objects than re-fixations on known objects.

Taken together, these findings suggest that when SA needs to be reacquired, participants quickly scan a diverse group of objects, with an emphasis on re-activating the previous goal-state. Yet when SA did not need to be reacquired, the goal state was intact so there was a decreased need to look at objects, a decreased need to look at objects quickly, and an increased need to detect novel changes in the environment. These results support the MFG Model (Altmann & Trafton, 2002) and the integrated framework for maintaining and recovering SA (St. John & Smallman, 2008).

This experiment focused on brief task-related breaks to force SA reacquisition, but there are many reasons that SA could become degraded over time and need to be reacquired. For example, an environmental interruption, a bathroom break, or a conversation with a friend could all lead to a need for SA to be reacquired. It is not completely clear whether SA after these breaks is "bad" or merely "not as good as it was before," but it is clear that poor SA leads to increased error rates (Jones & Endsley, 1996) and good operators need to reacquire any lost SA to remain effective.

A system that can identify the operator's cognitive state in real-time, such as their SA, and in response provide increased automation, or a cue, can result in improvements to task performance. Such a system will help fulfill the goal of the Office of the Secretary of Defense to get a single pilot to successfully control multiple UAVs simultaneously, and in addition, can result in improved performance on other dynamic tasks. While there has been some success in providing real-time feedback to an operator by predicting when they will make a post-completion error (Ratwani, Trafton, *in press*); a real-time cuing method designed to improve SA has yet to be successfully implemented, though attempts have been made (Trafton, Ratwani, & McCurry, 2010). Trafton et al.'s (2010) model indirectly measured SA based on task performance. These findings inform a systemic approach to evaluate SA non-invasively and in real-time.

ACKNOWLEDGEMENT

This work was supported in part by the Office of Naval Research under funding document N0001409WX20173 and N0001410WX30037 to JGT. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Navy. The authors thank the HAL lab and Missy Cummings for the use of the RESCHU simulation.

REFERENCES

Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review, 115*(3), 602-639.

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science, 26*, 39-83.

Altmann, E. M., & Trafton, J. G. (2004). Task interruption: Resumption lag and the role of cues. *Proceedings of the 26th annual conference of the Cognitive Science Society*, 42-47.

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin and Review, 14*, 1079–1084.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036-1060

Boussemart, Y. & Cummings, M. L. (2008). Behavioral recognition and prediction of an operator supervising multiple heterogeneous unmanned vehicles. *Humans operating unmanned systems*.

Crandall, J. W., Goodrich, M. A., Olsen, D. R., & Nielsen, C. W. (2005). Validating Human-Robot Interaction Schemes in Multi-Tasking Environments. *IEEE Transactions on Systems, Man, and Cybernetics Part-A, 35*(4), 438-449.

Durso, F. T., Truitt, T. R., Hackworth, C., Crutchfield, J., Nikolic, D., Moertl, P., et al. (1995). Expertise and chess: A pilot study comparing situation awareness methodologies. In D. J. Garland & M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness* (295–304). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37,* 65–84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 65-85.

Hodgetts, H. M., & Jones, D. M. (2006a). Contextual cues aid recovery from interruption: The role of associative activation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32*, 1120-1132.

Jones, D. G. & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine. 67*, 507-512.

Monk, C. A., Boehm-Davis, D. A., & Trafton, J. G. (2004). Recover- ing from interruptions: Implications for driver distraction research. *Human Factors, 46*, 650-663.

Office of the Secretary of Defense (OSD). 2005. Unmanned aircraft systems (UAS) roadmap, 2005–2030. Washington: DoD.

Ratwani, R. M., McCurry, J. M., & Trafton, J. G. (2008). Predicting postcompletion errors using eye movements. *CHI.*

Ratwani, R.M., McCurry, J.M., & Trafton, J.G. (2010). Single Operator, Multiple Robots: An Eye Movement Based Theoretic Model of Operator Situation Awareness. *Proceedings of the Fifth ACM/IEEE International Conference on Human-Robot Interaction.*

Ratwani, R.M. & Trafton, J. G. (in press). A Real-Time Eye Tracking System for Predicting Postcompletion Errors. *Human Computer Interaction*.

Salmon, P., Stanton, N., Walker G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics 37*, 225-238.

St. John, M., Smallman, H.S. (2008). Staying Up to Speed: Four Design Principles for Maintaining and Recovering Situation Awareness. *Journal of Cognitive Engineering and Decision Making, 2*(2), 118-139.

Trafton, J.G., Altmann, E.M. & Ratwani, R.M. (2009). A Memory for Goals Model of Sequence Errors. *Proceedings of the Ninth International Conference on Cognitive Modeling*.

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human–Computer Studies, 58*, 583-603.