THE IMPACT OF AUTOMATED PLANNING AIDS ON SITUATIONAL AWARENESS, WORKLOAD, AND SITUATION ASSESSMENT IN THE MONITORING OF UNCREWED VEHICLES

by

Grace Barnhart

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2023

Dalhousie University is located in Mi'kma'ki, the
ancestral and unceded territory of the Mi'kmaq.
We are all Treaty people.

Table of Contents

List of Tables

List of Figures

Abstract

Uncrewed vehicles (UVs) are pervasive in civilian and military operations. While technological advances allow most UV functions to be automated, human operators are frequently tasked to monitor UVs and intervene when system capabilities can not meet situational demands – often occurring as unanticipated situations. When monitoring UVs, good situational awareness (SA) allows human operators to recognize the need for their intervention. Although increasingly automated systems limit the need for human intervention, operator SA may be compromised as task functions are allocated to a system resulting in poor monitoring performance. In UV operations, the functions of route generation and selection may be critical opportunities for operators to build situational awareness. If a human operator, rather than a system, completes the processes encompassed in these functions (e.g., the evaluation the environment and comparison of route alternatives), the task considerations evaluated by the operator might bolster their SA and allow them to anticipate when interventions will be required to support the UVs. Improved SA is expected to result in better monitoring behaviours, fewer accidents, and improved overall performance. To evaluate the role that route generation and selection functions play in the monitoring of UVs, participants completed a virtual UV monitoring task with assistance from an automated route planning system. The system operated at three levels of automation (LOA) where the generation and selection functions were completed by either the operator human or an automated system – at the low LOA the system was an interface for the participant to generate the route, at the moderate LOA it generated routes which were ranked by the participant, and at the high LOA the system generated and selected one route. Once a route was determined, participants monitored the UVs as they traveled their outlined routes and intervened if the UVs were approaching an unsafe area. SA, situation assessment (visual attention), perceived workload, and performance were evaluated between the three levels of the automated system. The level of the automated system did not impact any variable that was assessed, indicating that the automation of UV route planning functions may not be detrimental to operator SA or overall performance. Findings and limitations can inform future examinations into the relationship between the planning process and SA, as well as the allocation of functions to a human operator or automated system to optimize performance in UV operations.

List of Abbreviations and Symbols Used

CaMP – Cognitive and Motor Performance

CI – Critical interval

COA – Course of action

CRACCEN – Command Reconnaissance Area Coordination and Control Environmental Network

GB – Gigabyte

IDEaS – Innovation for Defence Excellence and Security

LOA – Level of automation

NASA-TLX – National Aeronautics and Space Administration Task Load Index

RAM – Random access memory

REB – Research ethics Board

ROI – Region of interest

SA – Situational awareness

SAGAT – Situational Awareness Global Assessment Technique

SART – Situational Awareness Rating Technique

SPAM – Situation Present Assessment Method

UV – Uncrewed vehicle

Acknowledgements

First and foremost, I would like to express my deepest appreciation for my supervisor Dr. Heather Neyedli. I admire your ability to empower students by supporting their autonomy while maintaining open communication, providing guidance, and ensuring that fun is had along the way. I am also thankful to Finn (the dog) for the entertainment he has provided during our online meetings. To my supervisory committee members Dr. Aren Hunter and Dr. David Westwood, as well as my external reader Dr. Diane MacKenzie, I would like to extend a sincere thank you for the feedback, insights, and expertise you have offered throughout this process. I am incredibly grateful to have had the opportunity to learn from each of you.

I would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Nova Scotia Graduate Scholarship program, Innovation for Defence Excellence and Security (IDEaS), as well as Defence Research and Development Canada (DRDC).

A special thank you to Chelsey Hall, Brett Feltmate, Chris Holland and other members of the CaMP lab for their technological and morale support. Advice and anecdotes from your own research experiences have been invaluable. I am also grateful for Marcus Johnson from the SR Research support team for kindly answering my many, many questions. You were a life saver.

Finally, I would like to thank my family and friends for their support through the highs and the lows. I am especially grateful for my partner Matt. To say your support has been consistent throughout this program would be an understatement.

Chapter 1: Introduction

Uncrewed vehicles (UVs) extend operational limits by enabling safe access to hazardous or remote environments that would otherwise put human life at risk, and performing tasks with greater efficiency and endurance than can be expected of a human. UVs have become pervasive in both civilian and military operations, including search and rescue (de Alcantara Andrade et al., 2019), border surveillance and control (Haddal & Gertler, 2010), disaster management (Alawad et al., 2023), infrastructure inspection (Shakhatreh et al., 2019), reconnaissance (Schneider & Wildermuth, 2011), precision agriculture (Velusamy et al., 2022), climate change monitoring (Yuan et al., 2023), and exploration of other planets (Balaram et al., 2021). Across these operations, UVs can be equipped with varying sensors and deployed on the ground (Ni et al., 2021), over the water surface (Bai et al., 2022), underwater (Gafurov & Klochkov, 2015), and in the air (Laghari et al., 2023) to meet operation goals either through the collection of information or the performance of an action. While the capabilities, goals, and processes of UVs vary by domain, a task that is common to all operations is the determination of a route to the goal location. While a human operator typically determines a goal location for a UV, the routes traveled to these locations can either be controlled by an operator in real-time; prescribed by the operator and transmitted to the UV before deployment, and subsequently traveled semi-autonomously by the UV; or, generated and implemented autonomously by the UV itself. The route planning method used is often determined by either the technological capabilities of the systems or by the task goals. Increasingly, UVs have been implemented in swarms, where two or more UVs work together to achieve a common goal (Chen et al., 2020). To meet the complex computations necessary to coordinate the routing of numerous UVs, advanced automated routing protocols are being developed (Chen et al., 2020), where the route traveled to the end location is generated, selected, and implemented by an automated system within the UV.

Highly autonomous systems have minimal requirements for human input and excel in complex computations and repetitive actions. However, automated functions remain limited by their programming and the imagination of their programmers to anticipate future requirements - systems are ill-suited to adapt to situations that were not accounted for in their programming. Poor flexibility to unanticipated events is problematic in dynamic environments, such as warfare or search and rescue operations. Humans, on the other hand, generally have good contextual understanding and are better suited for decision-making in unpredictable situations (Hollister et al., 2017). Therefore, regardless of autonomy level, in dynamic environments UVs are typically monitored by a human operator that is expected to intervene in unanticipated situations where system constraints hinder performance. For example, if a UV's sensors do not detect an unanticipated hazard in the environment, an operator might recognize this and intervene to alter the UVs course to avoid the hazard.

Monitoring Uncrewed Vehicles

While employing a human to monitor UVs aims to maintain overall performance in situations outside the vehicle's capabilities, human operators frequently fail to recognize abnormal situations and the need to intervene. To recognize deviations from a normal situation, one must be aware of the current situation and anticipate what is expected to occur next under normal circumstances. Situational awareness (SA) encompasses an individual's perception and understanding of the contextual significance of stimuli in a given environment to know the state of that environment at a given time (Endsley, 1995a; Parasuraman et al., 2008). SA is the basis for decision-making where poor SA in monitoring tasks is associated with difficulty recognizing when to intervene and taking inappropriate actions while intervening (Wickens et al., 2021). Conversely, when one is aware of the state of their operational environment and can anticipate upcoming events, they are expected to interact with the environment efficiently to mitigate risks as they are presented. In UV operations data on the system and environment are often relayed to operators (e.g., location map, velocity). Even with these details available, humans tend to be complacent when monitoring highly automated systems (Parasuraman & Manzey, 2010). Failure to attend to the information provided impairs SA and the ability to recognize abnormal situations. This introduces a paradox, where systems that are designed to support humans ultimately impair their ability to perform their role. Allocating all tasks that can be automated to a system and leaving human operators to perform the remaining task of monitoring, leaves the human unfamiliar with the task and situation, and sets them up to fail in their monitoring responsibilities (Endsley & Kiris, 1995). Instead, the allocation of tasks should be designed to engage humans with relevant information that contributes to their SA without exceeding their cognitive abilities. To do so, the role that each task of UV operations play in the development of situational awareness and overall performance should be established.

Task Allocation and Situational Awareness

Endsley and Kaber (1999) outlined four functions that can be allocated to a human or an automated system: generating – developing potential courses of action selecting – determining which course of action to implement, implementing – carrying out the selected course of action, and monitoring – ongoing evaluation of the system or task status. The allocation of each function to either a human or automated system was described to create a taxonomy of the level of automation (Endsley & Kaber, 1999). An increasing level of automation (LOA) indicates that the system is carrying out more functions and reducing the workload of a human operator.

In uncrewed vehicle operations, the implementation of routes is often allocated to automated systems due to constraints in continuous communications that limit remote control (Chen et al., 2020; Li et al., 2022; NASA, 2022, n.d.b). As discussed previously, monitoring is often allocated to humans. The allocation of route generation and selection functions are less established. A route planning system that operates at a low LOA would allocate the processes of considering relevant variables to generate courses of action (COAs) and evaluating options to select a route to the human. Generation and selection of a route require the acquisition and integration of task-relevant information, which may contribute to the development of SA.

Current situational awareness guides expectations and influences how an individual assesses their environment, which dictates the sensory information they receive and their ongoing SA. Initial SA gained in the planning processes may then be critical as the basis for ongoing awareness (Endsley, 2015). However, generating routes can be computationally expensive. For example, to visit ten waypoints there are 181,440 route options (Wiener et al., 2009). In dynamic environments where numerous considerations must be factored in, the generation and evaluation of all possible COAs far exceeds the abilities of humans. For example, routes planned in a maritime setting must consider the coastal topography (avoid shoals), traffic (avoid colliding with other vessels), weather (wind), currents and tides (account for the direction of water flow), legal requirements (follow international and local maritime laws) and port capabilities (ensure docking is permitted and resupply is available) (Cummings et al., 2010; Woxenius, 2021). Rather than evaluate each option, humans often employ heuristics (cognitive shortcuts and assumptions) (Tversky & Kahneman, 1974). Heuristics can save computational processing at the cost of less optimal decision-making (Endsley & Kaber, 1999; Wiener et al., 2009). In route planning, the result is often the acceptance of any route that meets some minimal threshold held by the planner. While routes generated and selected with this process often meet day-to-day needs, they may not be sufficient in complex tasks (Wiener et al., 2009).

Conversely, a high level of automation can exploit the computational capabilities of automated systems, increasing the efficacy and efficiency of information integration (Colebank, 2008; Hilario et al., 2019; Lunghi et al., 2019; Stoyanov et al., 2016; Yu & Wei, 2009) and to optimize routes. However, as discussed, this high level of automation limits the opportunity for humans to build situational awareness through active task engagement. An alternative is to partially allocate route generation and evaluation to a system that presents alternative COAs, optimized to different task priorities, and allow the human to evaluate the COAs against task goals and constraints to select a route for implementation. A version of this strategy is used by current route recommendation software including Google Maps and Apple Maps. This allocation of tasks relieves the operator of the computationally heavy route generation but necessitates they integrate the available information by considering the environments current state and anticipating future situations to select the preferred option, which may bolster SA without overloading the operator. To date, no known literature has examined whether the generation and selection functions of route planning contribute to situational awareness.

The allocation of tasks to either a human or a system is not always static; adaptable systems can vary their LOA in response to a given parameter of the task, such as human workload. In these systems, each function can be allocated to best support the human operator to meet task demands exclusive of the LOA employed for other functions. Adaptable systems have been explored to lessen the effects of a high LOA on SA (Parasuraman et al., 2009) by reducing LOA when workload is low to keep humans engaged in a task where possible but transitioning to a higher LOA when their workload exceeds their abilities. However, it is unclear whether gains in SA acquired when the LOA is low extend to later occasions when the LOA has increased.

Naval Applications

Interoperability between humans and automated systems, and the effects of automated systems on situational awareness and decision-making is of particular concern to military operations. Defense Research and Development Canada (DRDC) is developing a computer-based network, known as the Command Reconnaissance Area Coordination and Control Environmental Network (CRACCEN) to aid with the collection, passage, and analysis of information onboard Canadian naval vessels (Tollefsen et al., 2017). Currently, DRDC is conducting research to understand the characteristics that are desirable in this automated command-and-control system. With the prevalence of UVs in all warfare environments, a subset of the network is anticipated to assist with the deployment and management of UVs. Detriments to situational awareness are of particular concern where time-sensitive decisions may impact human life, as well as national and international security. Further, federal directives require a human to be involved in the decision process of military operations (Banbury et al., 2022), emphasizing the need for autonomous systems to support human decision-makers.

The literature on human centered automation has a paucity of basic research to examine trends and fundamental effects outside of applied tasks. Studies examining the use of automated systems in route planning have been primarily conducted in naturalistic simulations with significant variations in task demands, variable measures, examined LOAs, and design between studies (Cummings et al., 2010, 2012). Coupled with contradictory findings and a lack of replication, the application of the current literature to UV monitoring is limited, and it is unclear how the automation level of a route planning system may impact human capacity to monitor the UV. Since naturalistic, high-fidelity simulations are often resource intensive, examining the effects of LOA at the basic level may be beneficial to further establish theories before progressing to applied research. Developing a basic simulation that can be shared and readily implemented without specialized equipment necessary for high-fidelity simulations would be beneficial to the development of literature in this area.

Purpose and Objectives

The aim of the present study was to evaluate whether the route generation and selection functions of the route planning process contributes to an operator's ability to monitor uncrewed vehicles. The primary objective was to evaluate situational awareness, situation assessment, perceived workload, and performance when these functions were allocated to either a human operator or an automated system. If engagement in these functions of the route planning process benefited the human's ability to monitor UVs, a secondary objective was to determine whether the benefits would remain if the planning functions were later allocated to a system. Although this research was conceptualized to contribute to the development of CRACCEN at DRDC and research on mixed-initiative systems in the Autonomous Systems Laboratory at the University of Waterloo, results are anticipated to inform future investigations on the allocation of tasks in any system where humans collaborate with uncrewed vehicles.

Outline

This dissertation consists of six chapters, where each chapter provides a basis for the following. Chapter one introduces the key concepts that pertain to the research question. Chapter two presents current theories and relevant literature to provide further context to the current state of knowledge and give rationale to the methodology. Once the literature has been presented, chapter two provides hypotheses relating to each objective. The methodology is outlined in chapter three to include the study design, the participant sample, and the measures examined. Chapter four reports the data analysis and results of statistical analyses. In Chapter 5, the discussion presents the results relative to their contribution to the greater literature, the limitations of the study and recommendations for future directions. Finally, the thesis is concluded in Chapter 6.

Chapter 2: Literature Review

Automated Systems

Automated systems are integrated into nearly all areas of human activities including health care (Manzey et al., 2011; Morrow et al., 2005), maritime navigation (Cummings et al., 2010; Loft et al., 2015), military operations (Ho et al., 2011; Knocton et al., 2023), and search and rescue operations (Al-Kaff et al., 2019). Across this wide range of applications, an automated system is any technology which selects data, transforms information, makes decisions, or controls processes (Lee & See, 2004). Systems may be designed to augment human performance or to complete tasks that humans cannot perform or perform poorly due to cognitive limitations (Wickens et al., 2021) such as limited working memory (Bouchacourt & Buschman, 2019), and computational speed (Blum & Vempala, 2020). Autonomous systems may also be introduced to reduce costs, limit human exposure to hazardous environments, or to meet productivity demands that exceed human abilities (Wickens et al., 2021). These systems can complete complex computations rapidly and excel at algorithm-based decision processes. A leading benefit to deploying automated systems is a reduced operator workload (Balfe et al., 2018; Endsley & Kaber, 1999; Tatasciore et al., 2020; Manzey et al., 2012).

Levels of Automation

Parasuraman et al. (2000) described the potential applications of automated systems across four task functions: information acquisition, information analysis, decision and action selection, and action implementation. Across these functions, the tasks allocated to an automated system vary along a continuum from manual performance (no automation) to completely autonomous (minimal human involvement). The level of automation describes where, between these extremes, a system operates (Sheridan et al., 1978; Wickens et al., 2021). Information acquisition is the process of retrieving data about the current state of the environment in which one is operating. In UV operations, a low LOA might collect and relay raw information such as coordinates, speed, and altitude to an operator. A high level of automation may alert an operator only if these parameters divert from expected values. At the information analysis function, a system with a low LOA may process data and present limited interpretations to the operator, such as mapping the movements of nearby vehicles and their projected trajectory. A highly automated information analysis system may integrate multiple variables to outline safe areas for a given UV to travel within. When determining which COA to implement, a decision and action selection system at a low LOA could provide a list of all possible COAs to a human operator, while a system at a high LOA might select a final decision without input from the human (Wickens et al., 2021). Finally, action implementation is the physical process in which the selected decision is carried out. At this stage, a low LOA might be a UV that is remotely managed by an operator, which might entail the control of propulsion systems and sensor positions. An example of a system operating at a high LOA could be a UV that navigates to an endpoint without operator input required.

Kaber and Endsley (1999) also outlined the potential functions of automated systems. Their functions are generating, selecting, implementing, and monitoring. While these functions describe similar tasks as those outlined by Parasuraman et al. (2000), the taxonomy of Kaber and Endsley specifies the function of generating COAs. At the generating function, a system with a low LOA might be a digital environment where an operator can map out potential routes, while a system with a high LOA might generate a single optimized route and present that to the operator. While the generation of COAs is presumed to be completed in Parasuraman et al.'s (2000) functions, it is unclear whether this would fall within the information analysis or decision selection functions. Since the present research is focused on the generation and selection of routes as discrete tasks, Kaber and Endsley's (1999) taxonomy is preferred here. The remaining functions are similar between taxonomy's – Kaber and Endlsey's selecting represents Parasuraman et al.'s decision and action selection, implementing refers to action implementation, and monitoring encompasses both information acquisition and analysis. Although a single system could perform at any combination of automation level for each function (Endsley & Garland, 2000; Parasuraman et al., 2000), there are some technological and practical limitations. For example, in a time-sensitive task it would not make sense to design a system with a low LOA that gathers and displays large datasets without the capabilities to integrate the data with a high LOA at a later function. The automation level of system should be determined by both task demands and system capabilities.

The allocation of tasks may be static between a human operator and a system or may change in response to real-time task demands. Adaptable automation is a system in which tasks can be allocated at varying LOAs as required throughout operations. Changes to a systems LOA can be initiated by the user, by the system itself under specific parameters, or through a cooperative process (Inagaki, 2003; Wickens et al., 2021).

Designing Function Allocation

When determining the automation level of a system, implications of allocating specific functions to the automated system must be considered across the larger system within which it operates. As an automated system assumes more task functions, operators transition physical and cognitive resources towards functions that cannot be automated yet or functions that are more suited to human capabilities, such as decision-making. The reallocation of cognitive resources from automated tasks increases under high workload or when multi-tasking is required (Parasuraman & Manzey, 2010; Wickens et al., 2021) and may be an attempt to save resources (Wickens & Dixon, 2007) or minimize effort. When systems operate at a high LOA, and cognitive resources are freed, performance is anticipated to improve in both supported and unsupported tasks (Kaber & Endsley, 2004; Manzey et al., 2012).

Though a reduction in workload and improvements in performance are promising, this is not always the outcome. Often, the introduction of an automated system changes the operator's role, thereby changing the type of work completed. As technological capabilities advance and systems operate at higher automation levels, the typical role of an operator has shifted from active task involvement towards passive supervision of systems (Di Flumeri et al., 2019). As more task functions are allocated to an automated system, additional responsibilities of operating, monitoring, and troubleshooting the system are created and fall to the human (Di Flumeri et al., 2019; Stapel et al., 2019). Shifting cognitive resources from active task involvement to passive monitoring can negate benefits and introduce new errors (Parasuraman & Riley, 1997). Likewise, the implementation of a system to complete a physical task decreases the physical workload of a human but places additional mental workload on the operator responsible for monitoring the system (Parasuraman et al., 2000). Consequently, it should not be assumed that introducing automated systems reduce an operator's overall workload.

In addition to an increasing mental workload, highly automated systems may induce performance decrements in monitoring, or when the operator needs to retake manual control. When tasks are completed autonomously, operators are poor at detecting malfunctions compared to when under manual control, are less capable of maintaining manual abilities (Clark et al., 2017; Manzey et al., 2012), and have poor performance when they must intervene (Kaber & Endsley, 1997; Layton et al., 1994; Reichenbach et al., 2011). These decrements are described by an out-of-the-loop unfamiliarity phenomenon (Wickens, 1992). As an operator is removed from the loop of controlling a system, they tend to rely excessively on the system and become unfamiliar with both the task and operational environment (Parasuraman & Manzey, 2010).

Out-of-the-loop unfamiliarity may be exacerbated within highly reliable but imperfect systems. When an automated system performs well, all feedback suggests the operator should continue to rely on the system. A lack of feedback can induce over-reliance, and complacency, which result in poor detection of system errors (Parasuraman & Manzey, 2010; Rovira et al., 2007; Rovira & Parasuraman, 2010). Because most automation is not 100% reliable, errors or unanticipated situations that require human intervention will inevitably arise. On these occasions complacent operators are less likely to intervene proactively. When monitoring of the environment and detection of errors are poor, abnormal situations can evolve into catastrophic accidents before they are recognized by an operator.

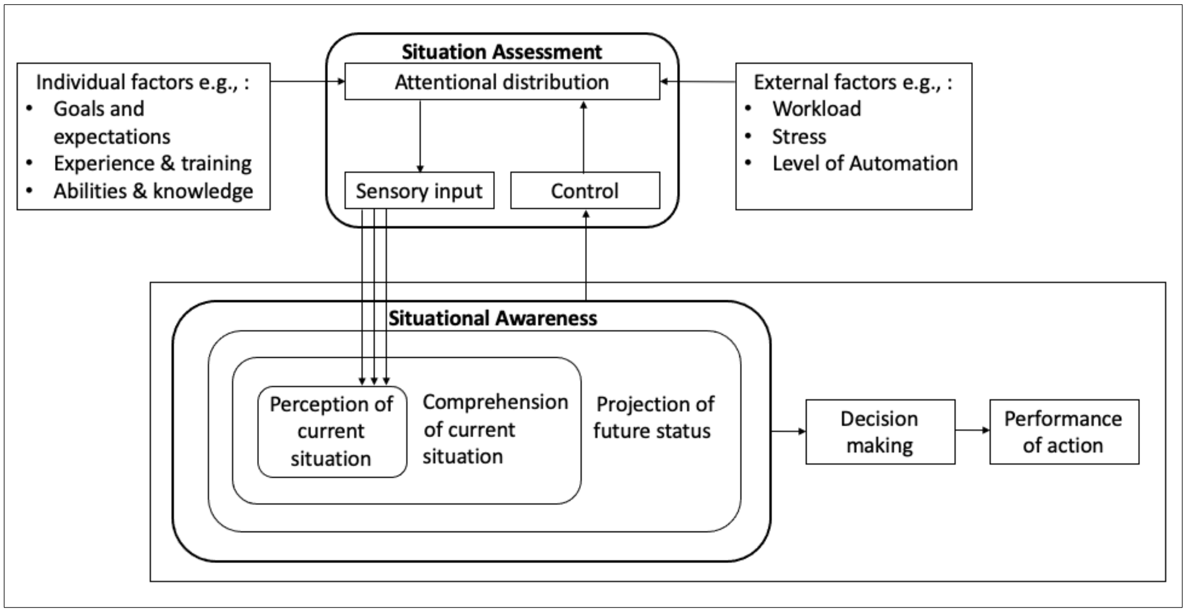
Though accidents are typically a result of many precipitating factors, rather than a single human error, the removal of a human operator from the control loop and the complacency that is induced limits human opportunity to recognize an impeding incident and intervene before they progress to accidents. Before implementing an autonomous system, the repercussions of system LOA on both operator and overall performance should be evaluated against the benefits of introducing the system. In determining the functions allocated to an autonomous system, considerations should include whether and to what degree workload will be impacted, whether the human is expected to retain manual skills or an understanding of the task, and implications if a system failure were to arise. Consideration should also be given to the appropriateness of tasks assigned to a human operator given the limits of cognitive resources and attention in humans.

Situational Awareness

Defining Situational Awareness

The relationships between workload (i.e., LOA) and performance in the same task may be modulated by an operator's awareness. Situational awareness (SA) characterizes the knowledge of the state of a dynamic environment. The technical definition of SA continues to be debated within the academic community due to differing applications between fields such as healthcare and engineering (Busby & Witucki-Brown, 2011; Parasuraman et al., 2008; Parse, 2018). While some describe SA as an insightful awakening (Parse, 2018), others view it as a measure of macro-cognitive awareness (Clark et al., 2017) or a mix of cognitive processing activities (Sarter & Woods, 1991). It is generally accepted that SA is neither a process nor a final product (Busby & Witucki-Brown, 2011) but a diagnostic state of an individual at a given time (Parasuraman et al., 2008). An end state of being situationally aware is never achieved: as the environment which one is aware of changes, their awareness must be updated. A prominent definition of SA in the field of human factors, and the one used to guide the present study, is *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* (Endsley, 1995a, p. 36). Within this definition, SA consists of three levels (1) the perception of stimuli, (2) understanding the meaning of the stimuli, and (3) the application of their meaning to anticipate subsequent events. Each of these levels builds upon the previous one; if someone does not perceive a stimulus in the environment, they cannot understand its meaning, nor can they project its implications to the future state of the environment. Endsley (2015) stresses that this hierarchical relationship of SA levels is not purely linear. When evaluating their environment, an individual's attention is biased by their expectations, experience, and knowledge, among other factors. Their biases dictate where their attention is directions and therefore the sensory input they receive. Sensory input in turn guides their perceptions of environmental cues, comprehension of the current situation, and their projection of how the situation will evolve (i.e., their SA). This first occurrence is linear; however, their initial SA becomes the basis in how their attention will be distributed in the future, which continues in a cycle (Figure 1). Those with good SA will correctly anticipate changes in their environment and direct their attention to the correct location at the correct moment to gather new information as it emerges. Correctly anticipating the location of critical information can facilitate the speed and accuracy in which information is evaluated (Endsley, 2015) and bolsters SA. From this cyclical relationship, good SA at the beginning of a task appears to promote the maintenance of SA throughout the task.

Figure 1. A model of the process in which situational awareness is developed and applied. Adapted from Hasanzadeh et al., 2018 and Endsley, 1995a.



While situation awareness refers to the state of knowledge (Endsley, 1995a), situation assessment is the means in which that state is achieved through the distribution of attention and sensory input (Figure 1; Endsley, 2015). Across many operational tasks, relevant information must be interpreted accurately and rapidly (Wiggins, 2021). Misinterpretations, or a failure to assess available information, have been noted to lead to errors such as a fatal accident that occurred off the coast of Brazil in 2009. On this occasion, an aircraft took off from Rio de Janeiro, Brazil and was flying over the Atlantic Ocean when it entered a tropical storm system and lost contact with air traffic control before crashing into the Atlantic waters killing all 228 people on board (De Wit & Moraes Cruz, 2019). Flight data and cockpit recordings recovered from the wreckage revealed that in the minutes preceding the crash some functions, including autopilot and airspeed data, were temporarily lost due to ice accumulation in tubes used to measure airflow speed. Rather than evaluating the remaining systems to understand the temporary nature and cause of the issue, the pilot inappropriately pulled up on the yoke, causing the aircraft to stall, which was the ultimate cause of the crash. An analysis of the events determined that the crash was due to human error and system design (De Wit & Moraes Cruz, 2019). Had the pilots assessed the available information they may have been aware of the nature of the technical issues and the flight may have ended differently. Automated systems can support efficient and accurate situation assessment, and decision-making, by drawing a user's attention to relevant information.

Situation assessment varies by task and context (Endsley, 2021). Through the direction of visual, auditory, and tactical senses one assesses their surroundings. A portion of this sensory information is then consciously perceived (SA level 1). Perceived information is incorporated with other information from working and long-term memory to recognize meaningful features (Blasch et al., 2006; Wickens et al., 2021), and to comprehend the significance of a stimulus (level 2) and its implication for the near future (level 3). In this process, the distribution of cognitive and sensory resources is influenced by situational awareness, because operators direct these resources to where they expect information to be. Situational awareness and assessment then occur in conjunction; situational assessment refers to how an individual physically interacts with environmental stimuli to receive sensory information whereas situational awareness is a cognitive state stemming from their interaction with the same sensory information.

Operationally, good situational assessment and awareness facilitate decision making (Endsley, 1995b; Hasanzadeh et al., 2018b; Lichacz, 2009). Though SA is the basis of decision-making (Endsley, 1995a) it is a unique process because all three levels of SA occur prior to the selection of a response (Wickens et al., 2021). Greater SA can improve response time and accuracy, particularly in response to sudden or unexpected events (Clark et al., 2017; Rovira et al., 2007), and may improve the detection of hazards (Hasanzadeh et al., 2018). When SA is low, improper decisions can be devastating. Recently, inappropriate SA was cited as a leading cause of two Boeing 737- Max 8 aircraft crashes (Endsley, 2019). In these two incidents, pilots were unfamiliar with updates to an onboard automated system, leading to incorrect comprehension and projection of the environmental stimuli. In both cases, the pilots took inappropriate actions under the updated system resulting in 346 deaths (Endsley, 2019). Levels two and three, which were affected in the accident described, have been identified as leading variables in the decision-making process (Chauvin et al., 2008), but level one SA remains vital as a means to higher levels. Each level of SA then contributes to an understanding of a situation at a given time, which informs decision-making.

Impact of Automation on Situational Awareness

When operators are less engaged in a control loop, or "out of the loop", due to a high LOA their SA is expected to be low (Endsley & Kiris, 1995; Smith & Jamieson, 2012). Conversely, systems at a low LOA have been linked to greater SA and performance in the same task but is also associated with operator fatigue and impaired performance in secondary tasks (Kaber & Endsley, 2004). It has been suggested that as LOA increases and humans are removed from a task, operators become passive recipients of information; cues may be perceived without processing their meaning or projecting their significance (Endsley, 2015) thereby reducing their SA. Others argue that decreases in information processing and SA correlated with highly automated systems are instead due to a conscious reallocation of attention (Parasuraman & Manzey, 2010). In the latter theory, operators may perceive task responsibility as being shifted to the autonomous system and, since it is not their responsibility, intentionally direct attention away from that task.

The notion of intentional cognitive strategies is supported by ecological research. Balfe, Sharples & Wilson (2018) studied railway operators (N = 21) while they monitored an autonomous railway system and intervened as necessary. The researchers suggested that rather than operators slipping into a passive state due to the automated system, they intentionally switched between active and passive monitoring according to task demands. Active monitoring behaviours (indicated by an upright body position) were used when operators expected to intervene. When the demand to intervene was lower, passive monitoring (when the operator leaned back in their chair) was more prevalent. Since active monitoring is a high workload activity (Warm et al., 1996), researchers ventured that shifts to passive monitoring are a strategy to maintain awareness of the environment while reducing workload (Balfe et al., 2018). Whether through an intentional strategy or a passive result of an autonomous teammate, operators direct fewer cognitive resources to the processing of cues during an automated task.

The effects of system LOA on operators have been examined in numerous tasks including remote lunar route planning between waypoints on topographic map interfaces (Cummings et al., 2012), railway traffic monitoring operations (Balfe et al., 2018), display aids on the bridge during a maritime vessel navigation task (planning and executing the task through a simulated waterway) (Nilsson et al., 2009), air traffic control (Metzger & Parasuraman, 2005), life support systems monitoring (Manzey et al., 2012) and on-road driving of autonomous cars (Stapel et al., 2019). While research across these tasks has contributed to concepts of task allocation, task demands as well as capabilities of human users and automated systems in each of these fields vary considerably (Kaber & Endsley, 1997). Consequently, recommendations of task allocation from these activities are not necessarily suited for UV operations.

Across all domains, when a task is entirely autonomous, the operator has less opportunity to develop and maintain their knowledge of the task (Manzey et al., 2012), and to create accurate mental models (Mogford, 1997). When systems are reliable but imperfect, the LOA at the decision selection stages may be critical to operators' ability to detect abnormal conditions (Rovira et al., 2007). When, by system design, the operator is not required to engage in the decision-making process, their cognitive processing to explore or place value on alternatives beyond those presented by the system is limited, which may reduce their SA and lead to a greater performance cost when systems err (Rovira et al., 2007). This suggestion is somewhat supported in a study of pilots (N = 30) completing a simulated enroute flightpath planning with a system operating at either a low (evaluated routes created by participants) or high (outlined a recommended course of action) LOA (Layton et al., 1994). Participants using the high LOA system took less time and used less effort in selecting a route than those using the system with a lower LOA but tended to rely on suboptimal courses suggested by the automated system (Layton et al., 1994). Though this study did not explicitly evaluate the act of assigning value to alternatives, there is increasing evidence to suggest that the process underlying all decision-making tasks is the assigning of value to each option and comparing alternatives, with a final choice aligning with whichever option has the greatest value (Merritt et al., 2015; Vaidya & Badre, 2020; cf. Williams et al., 2023). Good comprehension and projection of elements onto their environment, achieved through the development of accurate mental models in early decision-making processes, may then inform the subconscious assigning of value improving subsequent decision-making.

Adaptable automation may be helpful to reducing out-of-the-loop decrements associated with a high LOA. Adaptable systems may reduce mental workload compared to manual or static automated systems, while maintaining situational awareness (Parasuraman et al., 2009) because operators have some opportunity to engage in task functions. However, adaptable systems are not the entire answer to maintaining SA. Tasks are typically allocated to the autonomous system when the workload is high. Since the emergence of unexpected situations often leads to a high workload, adaptable systems may increase their LOA when automation capabilities are the most limited (Balfe et al., 2018). Additionally, the determination of whether to transfer a function to a system either creates additional work for the operator, or if allocated to the system itself, could result in unexpected transitions from autonomous to manual control, which has been noted to reduce SA compared to a planned take-over (Clark et al., 2017).

Nonetheless, performance improvements acquired in manual monitoring (under a low workload) have been found to persist when the system returns to a high LOA (Parasuraman, R., Mouloua, & M., Hillburn, 1999). Adaptable systems, where less tasks are automated (low LOA) while workload is low, may then improve performance in later highly autonomous periods while workload is high. It is unclear whether this is due to experience in processing task information or a change in monitoring strategy, nor whether the same benefits extend to different functions of the same task. Evaluation of whether periods of low LOA in route planning promote SA and performance in monitoring, and whether any gains persist when LOA increases would be valuable to inform the applications of adaptable automation.

Measuring Situational Awareness

Situational awareness is primarily measured through query-based evaluations that ask questions about the task at hand and compare responses to the known environment (See Endsley, 2021 for discussion and review). Leading techniques that used this approach are the Situational Awareness Global Assessment Technique (SAGAT) (Endsley, 1993) and the Situation Present Assessment Method (SPAM) (Durso et al., 1999; Endsley, 2021). The SAGAT uses a freeze probe method, where a task is randomly paused with the environment occluded, and participants are then asked a series of questions relating to the three levels of SA (Endsley, 1995b). SA is evaluated by whether responses are correct or not. This method is the most common and validated measure of SA (De Winter et al., 2019; Loft et al., 2015). However, the technique requires task interruptions and the need for participants to rely on memory while answering questions, which have been criticized for its limited applicability in naturalistic settings (Banbury & Tremblay, 2004). The SPAM also queries participants about past, present and future states of the environment but in this approach, queries are presented without pausing the task or occluding the environment. Here, situational awareness is evaluated as the time to correctly respond to queries. While this allows assessment of SA without the need to memorize cues, SPAM still interrupts the task and has been associated with secondary task decrements (Pierce, 2012). A recent review indicated that while the SPAM and SAGAT similarly predicted performance, SAGAT was more sensitive and less reliant on working memory (Endsley, 2021). Endsley's findings suggest that SAGAT may be the most useful and diagnostic measure of SA among query-based techniques.

A second approach to evaluating situational awareness in visual tasks is to infer awareness from situation assessment by evaluating the allocation of visual attention. As ocular muscle movements direct the fovea - the portion of the eye that perceives detailed information - around a visual search area, the position of the fovea compared to the environment is thought to indicate where the individual is attending and therefore gaining information. Varying methods can be used to evaluate visual sampling (Duchowski, 2007). A leading method is video-based tracking, where near infrared light is used to create corneal and pupillary reflections captured through video (Cognolato et al., 2018). The recorded reflections are analyzed to determine the relative position of the eye to evaluate whether a region of interest (ROI) was sampled. Although numerous metrics can be extracted from this method, visual sampling is most frequently expressed through two measures: fixations and saccades. Fixations refer to retinal stabilizations over a ROI lasting 150-600 ms. Environmental cues are perceived in these relative stabilizations and make up 90% of viewing time (Irwin, 1992). When a ROI is fixated, it is assumed that participants are extracting information from the region. The literature suggests that fixation count and duration significantly predict situational awareness (Moore & Gugerty, 2010) and performance. In an evaluation of visual sampling during a monitoring task in critical periods, when cues indicating the need for operator response became available, a strong correlation between visual sampling and performance ($r = .78$) was found to exceed the correlation between SA query scores and performance ($r = .20$) (De Winter et al., 2019). This finding suggests that task-relevant visual sampling may better indicate an operator's ability to perceive and apply relevant cues in the task environment. In the period between fixations, the eyes move rapidly (10-100ms) to the next stabilization, known as a saccade. During a saccade, the individual perceives limited cues from the environment (Duchowski, 2007). However, the frequency and direction of saccades contribute to situation assessment as the visual strategies employed by participants. In dynamic environments, task-relevant ROIs are expected to be fixated on more frequently at the cost of sampling other ROIs (Horrey et al., 2006). However, sampling of a ROI when task critical information is not present allows the confirmation of the state of the ROI and evaluation of when additional information may become available to avoid missing high priority cues. A strategy used among operators with good SA may be to frequently fixate on a ROI that will provide information and refrain from excessively sampling the same area when important information is not available (De Winter et al., 2019). Operators with good SA are expected to prioritize their sampling to have more frequent fixations of ROIs at critical moments but continue to sample the entire environment. One study that explored visual sampling and SA evaluated construction workers as they navigated a construction site to avoid hazards found eye metrics to significantly differ between participants evaluated as having either high or low SA (Hasanzadeh et al., 2018). Those with lower SA used more of their time evaluating immediate elements in the environment but dwelled less on hazards, while their counterparts with greater SA tended to direct their gaze to the horizon and dwell longer on hazards ahead of them. Greater fixation times on hazards may have allowed them to gain more information about the nature of hazards resulting in greater situational awareness, or their awareness might have alerted them to focus on these hazards.

Although eye movements have been closely linked to visual attention (Corbetta et al., 1998; Esmaeili, 2017; Hasanzadeh et al., 2017, 2018) and cognitive processes (Duchowski, 2007; Irwin, 1992; Just & Carpenter, 1980) it is important to note that fixations do not necessarily indicate that someone is attending to that region. Salient cues may be visually sampled but not consciously perceived (Strayer et al., 2004), and information outside of a visual fixation may be attended; though covert shifts of attention are often followed by overt shifts of visual sampling (Hoffman & Subramaniam, 1995). The perception of sampled cues is also affected by SA. Expected cues are more effectively perceived while salient but unexpected stimuli are often missed – a phenomenon known as inattentional blindness (Jensen et al., 2011). In this phenomenon, a failure to consciously perceive the unexpected stimuli, even when they are visually sampled, is due to a lack of cognitive processing rather than visual deficits. Since visual sampling does not reveal the knowledge held by an individual but rather an allocation of resources, eye tracking may be more appropriately viewed as situation assessment, a component of the process in which SA is gained.

As technology advances, there may be a greater opportunity to integrate physiological measures of operator state within human-automation teams. Eye-tracking measures overcome the concerns of working-memory and task interruptions that are associated with SA query-based assessments. Additionally, eye metrics gather continuous data allowing a more detailed analysis of participant response in both laboratory and ecological settings (Hasanzadeh et al., 2018). Evaluation of situational awareness and situation assessment in parallel can be used to determine both how individuals interact with their environment, and whether sensory input is appropriately perceived, understood and applied. Including performance measures can outline whether awareness was combined with procedural knowledge and skill to select and implement appropriate actions. Together, these measures can provide insight into the underlying aspects of task that are contributing to overall performance, and how these aspects might be impacts by automated systems.

Conclusion

Route planning and monitoring are essential components of UV operations that often involve uncertainty of dynamic environments and competing task demands (Cummings et al., 2010, 2012). Data integration for route generation is likely to exceed the cognitive limitations of working memory (Bouchacourt & Buschman, 2019) and computational speed (Fitts, 1951) of humans and can be allocated to automate systems. The allocation of route generation to automated systems is expected to reduce workload and allow attention to be directed to concurrent task demands. However, systems operating at a high LOA may also reduce operator SA and thus their ability to make informed decisions, particularly in response to unexpected situations (Manzey et al., 2012) compromising overall task performance. The selection of a route from options that have been generated by a system might be a critical function that encourages operator engagement and awareness without requiring excessive workload. However, it is unclear whether, and to what extent, the level of an automated system in the planning phase, consisting of route generation and selection, may affect subsequent monitoring of UVs.

Hypotheses

Recall that the purpose of this study was to determine whether engagement in the route planning process contributes to an operator's ability to monitor uncrewed vehicles.

The primary objective was to determine whether SA, situation assessment, perceived workload, and performance varied when route generation and selection were assigned to either a human operator or an automated system. To this end, participants completed a simulation where they monitored virtual UVs. In each trial the UVs route was planned, and then the UV was monitored as it traveled the set route through a dynamic environment.-Simultaneous to the UV monitoring, participants completed a search task. At the first block, all participants were assisted by a highly automated (high LOA) planning system that generated and selected a single route for each UV. In the second block, participants were assigned to one of three groups where the planning system operated at different LOAs. One group was assisted by a system with a low LOA, where the participant generated UV routes themselves, a second group was assisted by a system operating at a moderate LOA, where they evaluated routes generated by the system to mimic the route selection process, and the third (control) group continued with the highly automated system (high LOA). In the third and final block, all groups were assisted by the highly automated system.

Since all groups were assisted by the planning system operating at a high LOA in the first block, no difference was expected between the groups here. Similarly, no difference was expected between blocks for the high LOA group that exclusively used the planning system that operated at a high LOA. When participants were asked to generate (low LOA) or select (moderate LOA) routes in the second block, their perceived workload was expected to increase proportionally to the systems LOA. The low LOA group was expected to report the highest workload, followed by the moderate LOA group, and then the high LOA group. If engagement in the route generation and selection processes did contribute to one's ability to monitor UVs, the increased workload was expected to increase SA and performance. Therefore, it was hypothesized that in block two SA and performance would be inversely related to the LOA of the planning system – the high LOA group was expected to remain steady while the moderate and low LOA groups were expected to improve, with larger improvements in the low LOA group.

Situation assessment was evaluated through visual sampling metrics. It was hypothesized that if SA improved in the second block as anticipated for the low and moderate LOA groups, these groups would direct their attention to the UVs more often when new information was expected to be available. This shift in visual sampling would be expressed as more fixations when the UVs were nearing areas that could be unsafe, without excessively fixated on the UVs for the remainder of the trial, and therefore a lower proportion of overall fixations on the monitoring task. Though both experimental groups were expected to show these changes, larger improvements were expected for the low LOA group compared to the moderate LOA group. Saccades were evaluated without hypotheses there was little literature to suggest how they may be impacted. As with the other variables, situation assessment was expected to remain consistent between blocks one and two for the high LOA group.

If the allocation of planning functions did impact any of the examined variables, differences observed between groups in block two may stem from a greater familiarity of the specific UV routes that were generated (low LOA) and selected (moderate LOA) for each trial, or they may reflect broad improvements in the process of monitoring UVs, such as refined mental models of the task. A secondary objective was to explore whether effects of group would remain when planning functions were reallocated to the automated system in block three. This change in the systems LOA reflects adaptable automation. If a planning system operating at a low or moderate LOA results in broad and persistent improvements in monitoring, gains in SA, performance, and situation assessment in block two would be expected to remain when all groups return to the high LOA in block three. This was explored without specific directional hypotheses. Finally, in block three perceived workload is expected to return to initial levels for the low and moderate LOA groups and remain consistent for the high LOA group.

Chapter 3: Methodology

Participants

A university-based psychology participant pool, and word of mouth were used to recruit forty-five adults from the Dalhousie community. Data from one participant could not be extracted resulting in a total sample of forty-four. Participants ranged from 18 to 39 years of age (M = 21.23, SD = 3.96), with 72 identifying themselves (gender) as a woman (32), 25% as a man (11), and 2.3% as non-binary (1). The psychology pool was open to all students registered in a psychology course at Dalhousie University, regardless of their primary area of study, and was the primary method of recruitment. Students recruited through the psychology pool registered through an online portal and were awarded bonus points towards a psychology course for their participation. Word of mouth was only used if community members expressed interest in the study but were not registered in a psychology course.

An a priori power analysis indicated that a sample of thirty-three (eleven per experimental group) participants was necessary to achieve power of 0.8 to detect a moderate effect size of .65 for interactions (within and between-group; $\alpha = .05$, ß = .2). This was rounded up to fifteen per group to account for potential data loss due to incomplete sessions or technical issues. To be eligible for this study participants were required to: have normal or corrected to normal vision (i.e., 20/20); to be between the ages of 18 and 65; and, to have no known neurological or visual conditions that would restrict the coordination of their eye movements, their visual or cognitive processing abilities, or the fine motor control of their right upper limb. Normal vision was necessary to view and complete the tasks on the computer screen. Those with visual acuity below

20/20 were expected to squint more often to view the on-screen stimuli, which can be problematic for measures of visual sampling. For the same reason, participates were asked to avoid wearing eyeglasses where possible. Visual acuity was not measured within the study but was included to encourage self-screening for participants that were not able to see the screen well enough to complete the simulation. The age range corresponds to the adult working population in Canada. The exclusion of those with known visual or neurological conditions was necessary to ensure participants would be able to interact with the simulation. Fine motor control of the upper right limb with the ability to manipulate a computer mouse was necessary to complete the simulation.

Participants attended a single two-hour session at the Cognitive and Motor Performance (CaMP) lab at Dalhousie University. All participants were required to read and acknowledge an informed consent form (Appendix C) prior to participation and were given time to ask questions or voice concerns they may have had prior to taking part in the study. The informed consent specified that participants could withdraw from the study at any time during the session without limitations by communicating to the experimenter that they no longer wished to continue. Participants were asked to provide their demographic information of age and the gender with which they identify. There is no evidence to suggest that gender or sex influence the variables examined in this study, so age and gender were collected only to provide context regarding the sampled population. A participant ID code was used, so identifying information was separate from the experimental data throughout the study.

Participants who registered through the psychology pool received two credit points towards a psychology course. All participants were informed that they may receive an honorarium of up to 10 CAD if they perform well, and that performance would be based on the number of successful trials in the experimental blocks. To fairly compensate participants across experimental groups, everyone received 10 CAD regardless of their performance. The honorarium was offered to incentivize participants to perform well and has been used previously (Neyedli et al., 2011) to promote participant engagement in tasks throughout data collection. Although this technique requires deceiving participants, potential negative outcomes were limited as participants recruited through the psychology pool were guaranteed two credit points, and all participants received what they expected to be the maximum honorarium. The deception was well received by participants during the debrief.

Design

In general, UV operations include three components – planning the UV route, monitoring the UV as it travels the planned route, and evaluating data that is collected by the UV. To mimic these task demands, participants completed a computer-based simulation where each trial consisted of two phases. In the first phase, a route was determined for a UV in a dynamic environment (see *Apparatus* for further details). In the second phase, the UV traveled the route and participants monitored the UV as it traveled the outlined route. If the UV approached an area that was not safe for its passage, the participants were expected to intervene (see *Apparatus for further details*). While monitoring the UV, a concurrent search task was completed to mimic searching data collected by a UV (e.g., images, a video feed) while monitoring the UVs position.

A mixed design was employed to evaluate how the automation level of a route planning system impacts a human's ability to monitor a UV (phase two). LOA was manipulated between experimental groups and across the blocks. In the first of three experimental blocks, all participants were assisted by a route planning system operating at a high LOA. In the second block participants were randomly assigned to either continue with the system at a high LOA (control), or to systems that operated at either a low or moderate LOA. In the third experimental block, all participants were assisted by the system operating at a high LOA. As detailed in the *Apparatus* section, the sole difference between the experimental groups was the automation level of the planning system in block two.

Each block consisted of twenty-two trials and each trial included two UVs. In each block, two of the twenty-two trials were used solely to evaluated SA and were not included in the analysis of any remaining variables (see *Measures* for further details.) In five of the remaining twenty trials, participants were required to intervene with one of the UVs for a total of 12.5% of UVs requiring intervention. The remaining UVs (87.5%) would successfully complete their routes without intervention. Since complacency is more likely when automation is highly reliable (Parasuraman & Manzey, 2010), this rate of intervention was chosen to promote complacency while maintaining a relatively high number of intervention trials. Additionally, this ratio is well above the suggested 70% reliability for automation to be perceived as useful (Wickens & Dixon, 2007), and therefore maintains some ecological validity.

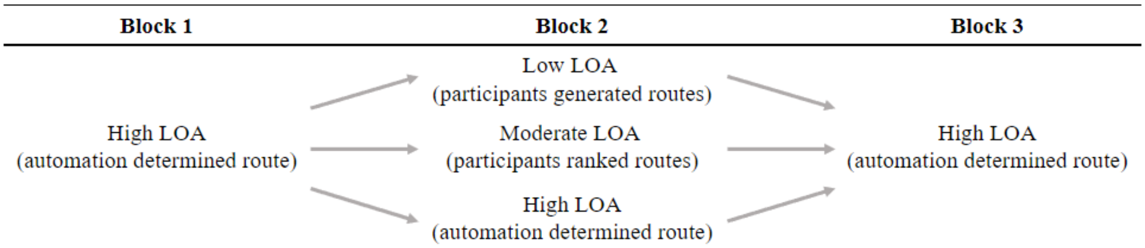


Figure 1. The automation level of the route planning system used for each group across the experimental blocks.

Apparatus and Measures

Apparatus

To measure visual sampling participants wore a head mounted EyeLink®II eye tracker (Figure 3). The EyeLink®II is a high-resolution video-based system that approaches the accuracy of the scleral search coil – the gold standard of eye movement measures (McCamy et al., 2015). The head-mounted configuration allows free head movement. Once the eye-tracker was mounted, a nine-point calibration and subsequent validation were completed. Data collection did not proceed until validation was achieved with a maximal point error no greater than 2.0°, and an average point error no greater than 1.5°. If calibration had to be repeated within session, the same standard was upheld. Measures were sampled at 250 Hz with pupil-corneal reflection (P-CR) tracking wherever possible for high acuity and low noise. If calibration was not achieved after numerous attempts with P-CR, pupil only tracking was used.

Figure 2. The head mounted configuration of the EyeLink®II system.

As indicated previously, participants completed a simulation that mimicked the demands of a non-descript UV operation - the UV route was planned (phase one) and then monitored (phase two) in a dynamic environment where data collected by the UV was examined concurrently to the monitoring of the UV. While this simulation was not designed to mimic any specific task, a theoretical example of this task is the responsibilities of a search and rescue operator that reviews images collected by a UV for signs of hazards or a missing person while monitoring the UVs position in the area to be searched.

When introduced to the task, participants were told they would be paired with UVs to explore a remote region (Appendix F). It was explained that portions of a larger region were represented in small grid segments (Figure 4). In each segment, some areas were known to be safe to travel through, some were known to be unsafe to travel through, and others were not known to be either safe or unsafe for travel. The status of each grid square was indicated by its color: white squares were safe, black squares were unsafe and should not be travelled through, and grey squares were not confirmed as either safe or unsafe. Since it was unknown whether the grey squares were safe, they presented a risk to the UVs and were therefore referred to as risky areas. The UVs current location was depicted by a white square outlined in blue, and a goal end point was indicated by a red rectangle (Figure 4). Participants were told that the goal of the simulation was to identify safe routes through the risky (grey) areas by having the UVs travel from their start point to the goal end point. Importantly, all possible routes between these points required the UVs to travel through one or more risky areas.

To travel between these points, the (fictitious and virtual) UVs required route instructions. To generate routes, participants were assisted by a (simulated) automated route planning system operating at one of three LOAs (detailed below). Participants were told that when the route was determined in the planning phase and sent to the UV, the UV would be able to follow the indicated route. It was explained that once the operator indicated that the UV could start, the UV would begin moving along the route and the monitoring phase would begin. Participants were also told that the UV was equipped with sensors that could determine whether the area directly in front of them, to a range of two grid squares, was safe. Risky areas that were determined to be safe by an approaching UV would be indicated by grey areas of their grid segment map updating to white. Areas that could not be determined as safe remained grey. The participant's role, as the operator of the UV, was to monitor the UV's movement on the grid segment, and if a UV was too close to a risky area, to stop their movement by clicking a "CANCEL" button. If all risky areas included in a UVs route were determined to be safe the UV would be able to pass from the start point to the goal end point without travelling through a risky area, and the segment was successfully completed if the participant did not intervene. If any risky area on a route remained grey, the segment was successfully completed if the UV's route was cancelled while the vehicle was within the grid square immediately preceding the risky area. Most routes were safe for the UV to travel, so the monitoring of a single vehicle was easy. To increase task demands on the participants, two UVs in unique segments were monitored concurrently as they followed their respective routes for their segment, in each trial (Figure 5).

To mimic the secondary task of evaluating data collected by a UV, the monitoring phase included a concurrent search task that started when the UVs initiate the movement along their routes. In the search task, participants located and clicked on a target "T" among an array of nine distractor "L's" (Figure 5). Regardless of whether, and how quickly, a response was entered, a new array to be searched was displayed every five seconds. If the target was clicked on within the five second window, the iteration was considered a success. A timer just above the search task counted down from five for each of these intervals, and iterations of the search task continued for the duration of the monitoring phase. The search task continued through five second iterations until both UVs stopped moving, marking the end of the monitoring phase. A UV could be stopped in one of two ways - either it reached the end point or was canceled by the operator. Pilot testing was conducted to determine the appropriate array size and time interval for the search task because previous studies have found that tasks that are too easy do not elicit the benefits of automated systems as the human is not cognitively taxed and may easily complete simultaneous tasks (Endsley & Kiris, 1995; Rovira et al., 2007). To determine the appropriate number of distractors, the serial self-terminating search (SSTS) model (Sternberg, 1966) was employed. The model indicates that search time is dependent on the time to search each object multiplied by the number of objects to be searched divided by two because, *on average*, a target will be found halfway through a search (Wickens et al., 2021). Pilot testing determined the average time to locate a target at a 90% success rate and used the SSTS to determine the appropriate number of distractors for a five second interval.

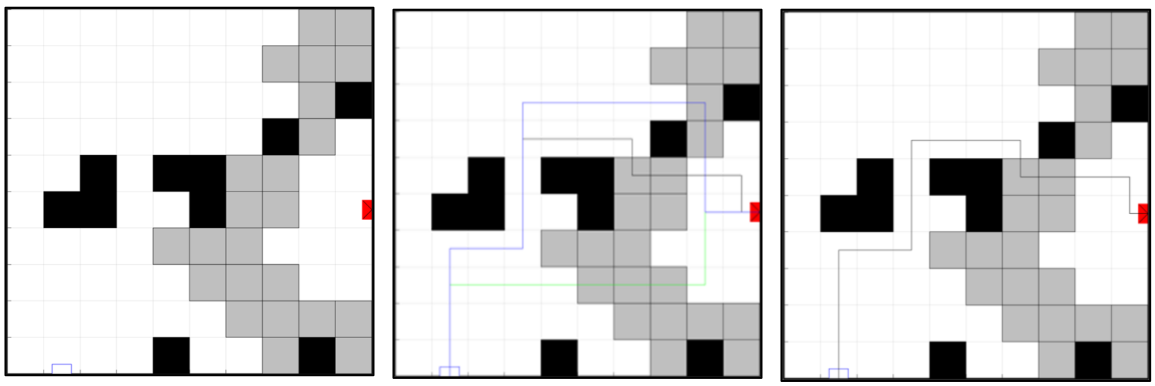


Figure 3. Examples of the route planning task at each level of the automated planning system. At a low LOA (left) the system prompted participants to generate three potential routes. At a moderate LOA (center) the system generated and displayed three routes (shown in black, blue, and green here) and prompted participants to rank the routes from their most preferred to least preferred. At a high LOA (right), the system generated and displayed a single route for the UV.

The automation levels of the route planning system were as follows. When the system operated at a high LOA, a single route - that minimized the number of risky areas included in the route and the overall route distance - was generated and transmitted it to the UV. Here participants were not involved in route generation or selection. At the moderate LOA, the system generated potential routes with varying total distances and risky areas included and asked the participant to rank the routes from best to worst based on their individual preferences to mimic a route selection process. In some cases, the generated routes traded-off between minimizing the total distance and the number of risky areas included, while in others the routes were equally optimal. Finally, at the low

LOA the system prompted participants to generate three route options by clicking on the grid to indicate successive waypoints. Here, participants were instructed to minimize both the total distance traveled and the number of risky areas included in the route. To maintain study duration and experimental control between groups, the final route was selected by the automated planning system for all trials. Therefore, the route generation and selection processes associated with the low and moderate automation level systems, respectively, were intended to simply engage participants in these functions of planning to determine whether the processes influence their SA.

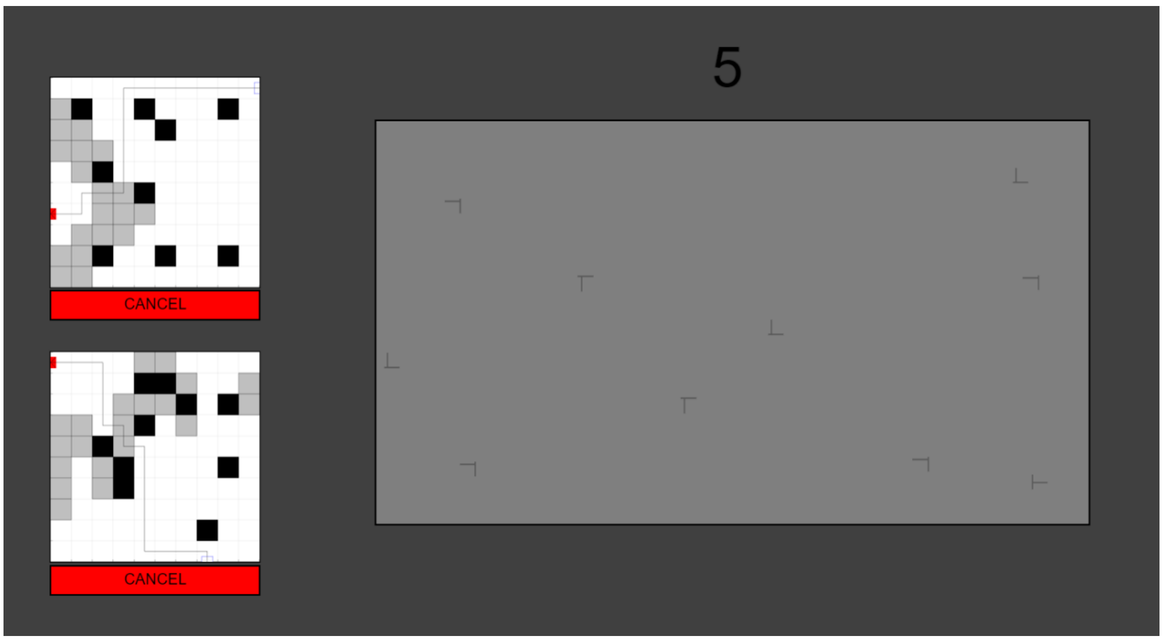


Figure 4. An example of the monitoring phase of the simulation. On the left, the uncrewed vehicles, represented as a white square with outline, simultaneously follow their respective routes to the red square. On the right, the array and target ("T") locations refresh every five seconds.

The simulation was displayed on a monitor (dim. 59.7cm x 33.6 cm, res. 2560 x 1440) situated at approximately 57 cm directly in front of participants. An identical display monitor positioned to left of the primary monitor and rotated towards the participant was connected to a computer that was separate from the EyeLink® system. This monitor was used for planning exercises, and measures gathered through online survey (i.e., workload and situational awareness). Participants interacted with each interface through mouse clicks. The simulation was programmed with Experiment Builder (v 2.3.527) and MATLAB (R 2017b). The grid segments and videos representing the travelling UVs were created in MATLAB and loaded into Experiment Builder as .mp4 files. The remaining components of the simulation were programmed in Experiment Builder.

Figure 5. The experimental set up. The primary monitor displays the simulation. Although the monitor on the left does not display anything while the simulation is running, planning exercises as well as SA and workload measures were completed on this adjacent monitor.

Measures

Situational Awareness

Since the participants goal for the UV monitoring task was to ensure the UVs did not enter the risky areas, knowing when the UVs were approaching a risky area was critical. The UV sensors had a range of two squares and would update the segment by the time the UV was within one grid square of the risky area. Awareness and sampling of the UVs in this interval was then of particular concern. The period in which the leading edge of the UV first became within two grid squares of a risky area, until the leading edge entered (or would have entered if the square updated to indicate it was safe) a risky area was noted as a critical interval (CI).

SA was assessed twice in each block, once when one of the UVs were within a CI, and once when neither were within a CI. SA was evaluated with the Situational Awareness Global Assessment Technique (SAGAT) an extensively validated measure used frequently in the field of human factors (Endsley, 2011, 2020; Endsley & Garland, 2000). As described in chapter two, the SAGAT entails freezing and occluding the simulation, and then asking participants questions about the state of the simulation immediately preceding the freeze. Answers are then compared to the true state of the simulation to determine participants' awareness. SA was measured as the proportion of queries successfully answered. The proportion for each block was then the total number of correct responses divided by the number of queries posed. Queries are designed to assess the three levels of situational awareness (perception, comprehension, and projection). For example, participants were presented with an image of the one segment, with the UV, the route, and the start/end points removed, and asked to click on the location where the UV had been. See appendix H for the full list of queries. Between iterations, queries varied in the segment referenced, denoted as either top or bottom to relay their position on screen. Queries that asked for input to identify the location of items within the simulation (questions one and two in appendix H) were presented with MATLAB, the remaining queries were presented in Opinio, an online survey platform.

Eye Movements

Visual samples were collected relative to three ROIs: the two segments that represented the UV environments and the visual search task each contained in their respective borders. The proportion of visual attention allocated to each ROI was evaluated as the proportion of dwell time spent within that trial. Sampling strategies were also assessed by examining the proportion of saccades that occurred between the search and monitoring task, which indicates the flow of visual sampling. Together, these measures describe the visual sampling strategies employed by participants under competing task demands. In addition to overall strategies, the proportion of UVs that were fixated while within their two square range of a risky region were evaluated to determine whether participants were sampling the UVs in this CI. The eye-tracker was worn for the entirety of the three experimental blocks, but measures were evaluated only in the monitoring phase of each trial.

Workload

Objective workload was manipulated between groups as the automation level of the planning system. When the system operated at a low or moderate LOA, participants had to complete additional steps to generate the routes.

Perceived workload was evaluated at the end of each experimental block. Participants completed a modified version of the National Aeronautics and Space Administration Task Load Index (NASA-TLX; Hart & Staveland, 1988) in which they rate their workload in six dimensions – mental demand, temporal demand, effort, physical demand, and performance (Appendix G). Dimension definitions were provided (Table 1). The NASA-TLX is a widespread measure of workload (Hart, 2006). Rather than the traditional 21 point-scale, each dimension was rated on a ten-point scale due to an error in converting the questionnaire to an online format. In an effort to facilitate experimentation and recruitment, by keeping the time commitment required for study participation within two hours, raw scores were used rather the traditional weighted score (see Bustamante & Spain, 2008; Hart, 2006 for a detailed discussion of modifications to the NASA-TLX). The workload questionnaire was modified in the performance dimension, so a greater score was associated with better performance rather than a greater workload – which was reversed prior to analysis to align with the original format.

Table 1. NASA-TLX dimension definitions (NASA, n.d.-a).

Scale Title	**Endpoints**	**Description**
MENTAL DEMAND	*Low/High*	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	*Low/High*	How much physical activity was required (e.g., pushing, pulling, turning, controlling, etc.)? Was the task demanding, slow or brisk, slack, or strenuous, restful or laborious?
TEMPORAL DEMAND	*Low/High*	How much time pressure did you fell due to the rate or pace at which the tasks or elements occurred? Was the pace slow and leisure or rapid and frantic?
EFFORT	*Low/High*	How hard did you have to work (mental and physical) to accomplish your level of performance?
PERFORMANCE	*Good/Poor*	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing this task?
FRUSTRATION LEVEL	*Low/High*	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Procedure

Once the informed consent process was complete, participants sat in front of a computer workstation for the duration of their study session. They were told they would act as an operator of UVs in a simulation exploring a remote region. Participants were introduced to the task progressively across three training blocks, and then completed three experimental blocks. Prior to each block, a script explaining the upcoming task (Appendix H) was displayed on screen. Once they had read each script, participants were given an opportunity to ask questions.

Participants were first trained on the two components of the simulation individually to allow them to become comfortable with each. In the first training block, they completed ten trials of the search task and were offered feedback by the researcher. The second training block consisted of ten trials. In each trial, participants monitored a single UV as it travelled through a unique segment. To first introduce them to the task process and environment, the first five trials did not require their intervention. To allow them to practice the timing of interventions, the remaining five trials required intervention to prevent the UV from entering a risky region.

To allow adaption to the entire experimental set-up in training, the EyeLink®II was donned between the second and third training blocks. A nine-point calibration and validated were completed, as indicated in the *Apparatus* section. The EyeLink®II was worn throughout the remainder of the study.

The final training block brought the two components together and consisted of five training trials that followed the format of an experimental trial. For the final training block and all experimental blocks, each trial began with an eye-tracker drift correction.

Next, two grid segments were displayed on the left of the screen (Figure 5) each outlining the status of the squares, the UVs current location, the end location, and the selected route. Participants could review the segments and selected route for up to one minute or click on a "START" button to proceed before the minute had elapsed. Once the review was completed, the monitoring phase began. At the onset of the monitoring phase the UVs travelled their outlines routes while the search task appeared on the right of the screen (Figure 5). The search task continued with five second iterations as long as at least one UV was moving. Once both UVs stopped moving, either upon reaching the end point or being cancelled by the participant, the trial was completed.

Once training was complete, participants were given the opportunity to take a self-paced break before moving on to the experimental blocks. In the first experimental block, each trial followed the sequence of drift correction, review the segment and routes generated by the planning system, and then monitor UVs while locating targets. SA data was measured at two random intervals within the block. As discussed in the *Apparatus* section, when SA was assessed, the simulation was paused and occluded, and then the participants were asked a series of questions regarding the state of the simulation immediately before it was paused. Once the SA queries were completed, the same trial was not resumed as the interruption of the task was expected to affect their performance. Apart from the SA measure, no data was collected or evaluated for trials that included SA assessments. The order of trials presented within each block was randomized for each participant. Following the block, participants completed a workload assessment, doffed the eye-tracker, and then took a self-paced break.

The second experimental block followed a similar procedure as the first, however, the automation level of the planning system was manipulated between the three experimental groups. At the beginning of each trial, participants were directed to the adjacent monitor to complete planning exercises for their respective LOA. For participants in the low LOA group, a segment that included the status of each grid square, the current UV location, and the end point, was displayed. They were asked to recommend three potential routes for the UV by clicking on the segment to indicate waypoints, which were joined on the display to represent their recommended route. Once the three routes were entered, the planning process was repeated for the second segment in that trial. Finally, they reviewed the route selected by the automated system, that the UVs would follow, for each segment and continued to the monitoring phase as in the first experimental block. Participants in the moderate LOA group were presented with a segment that outlined three potential routes and were asked to rank them from their preferred route to their least preferred, accounting for the overall length of the route, the number of risky areas included on the route, and personal preference. Once routes were ranked for the two segments in each trial, participants reviewed the autonomously selected routes that the UVs would follow and continued to the monitoring phase. Those in the high LOA group, were also directed to the adjacent monitor at the beginning of each trial where they reviewed the UVs routes individually for 20 seconds each to maintain the shifting of attention and time required for planning in the other groups. As done in the other groups, they were presented with the routes on the main monitor and then continued to the monitoring phase.

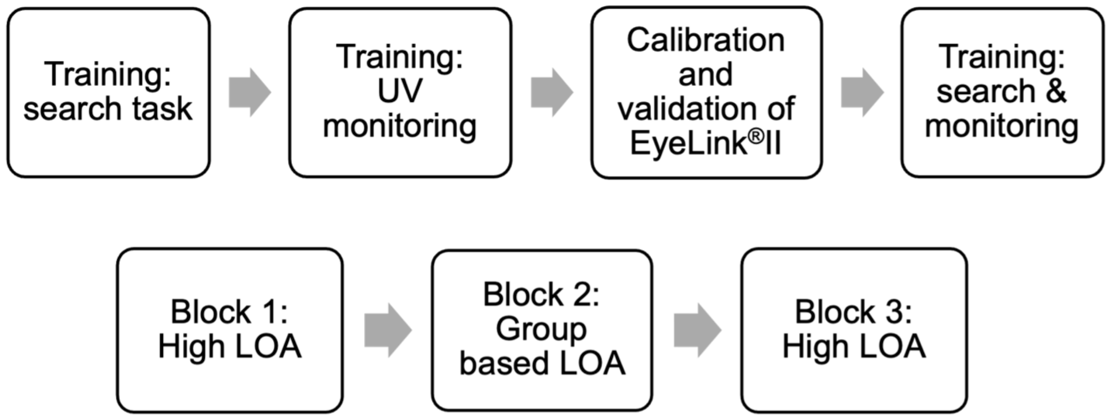


Figure 6. The study procedure. Situational awareness measures were gathered twice within each experimental block, and workload will be collected with a NASA TLX questionnaire following each of the experimental blocks.

To maintain experimental control, routes for a given trial were consistent for all participants and did not account for their input in the planning exercises. Pilot testing indicated that the planning tasks were intuitive. From this feedback, and since the researcher could respond to participants questions if required, and performance was not evaluated, training on the planning tasks was not included. All aspects aside from the planning exercise of each trial remained the same as in block one for all groups. SA was evaluated at two random intervals, and workload was assessed upon completion of the block before a self-paced break.

For the third and final experimental block, all groups used the highly automated planning system, and the twenty-two trials followed the same format as experimental block one. SA was evaluated twice, and workload was assessed upon completing of the block. After the third experimental block, participants were debriefed on the goals of the study, the experimental groups, the process of removing their data from the study or reporting concerns about the study, and the honorarium (Appendix I). Once debriefed, participants were again given the opportunity to ask questions, and received their honorarium.

Data Analysis

Workload

Data was extracted from Opinio, and further processed in MATLAB. Questions were reversed where necessary so increasing values were associated with an increasing workload. Scores were summed for each questionnaire creating a workload score for each block for each participant.

Situational Awareness

Data was extracted from Opinio and MATLAB, and further processed in MATLAB. Participant responses were compared to the true responses (coded as 1 = correct, 0 = incorrect). When examining the data, it was noted that some questions had a disproportionately low number of correct responses. In trial 27 participants were asked to indicate the start point of one of the UVs to determine their perception of cues in the simulation. Only 4 responses were correct, and the remaining responses primarily reflected other grid features (UV end point, UV location) indicating a poor understanding of the question rather than a poor perception of cues. Two questions in trial 55 asked participants to compare the UVs location relative to other task features when the UVs were equidistant, each question had only one participant enter the correct response of "not applicable". Therefore, these three questions were removed from analysis.

The two situational awareness assessments from each block were combined, and the total number of correct responses was divided by the number of questions to determine a proportion of correct responses for each participant at each block.

Performance

Data was extracted from Experiment Builder and further processed in MATLAB. Performance was not evaluated on trials where situational awareness was assessed because the simulation was paused and not re-started for the situational awareness assessment.

Monitoring Task. Segments were successful completed when a UV completed the route without entering a risky area or when a UV that approached risky areas was cancelled within one grid square of the area. The number of successfully completed segments was divided by the total number of segments monitored (i.e., two per trial) to determine a proportion of successfully completed segments for each participant at each block.

Search Task. The number of successfully located targets was divided by the number of arrays searched in each block to determine the proportion of successfully located targets for each participant at each block. In cases where a trial ended in the middle of a search iteration and no response had been entered for that array, the iteration was excluded from analysis.

Eye Metrics

Fixation and saccade data was extracted from Experiment Builder and further processed in MATLAB. There were three regions of interest (ROIs), defined as the area of each monitoring task as well as the area of the search task. Fixations were evaluated as the proportion of dwell time in each ROI in each block. Critical fixations were evaluated as the proportion of monitoring ROIs that were fixated as the UVs were with two units of a risky area. Saccades were evaluated as the movement between the ROIs regardless of whether other stimuli (e.g., the timer) were fixated between ROIs. Visual sampling was not evaluated on trials where situational awareness was assessed because the simulation was paused and not re-started for the situational awareness assessment.

Data loss was evaluated as the percentage of samples without data. First, the number of blinks were multiplied by the average blink duration in each trial to determine the duration where the participants gaze was not recognized by the eye tracker for that trial. This number was then divided by the duration of that trial to determine a proportion of data loss (SR Research, 2021). Mean data loss was calculated for each participant. On average, there was minimal data lost ($M = 1.3\%$, $SD = 1.9\%$, $MIN = .06\%$, $MAX = 9.0\%$,)

Prior to analysis, the distribution of the data for each variable was examined and outliers were investigated. Aside from the SA questions detailed previously in that section, no data was removed from analysis. A 3 (Group: low, moderate, high) by 3 (Block: B1, B2, B3) mixed ANOVA was completed with IBM SPSS Software (version 28.0.1.1) for each measure to determine whether the measure was impacted by the automation level of the planning system (between groups - primary objective) and whether effects lasted when the system returned to a high LOA (within groups – secondary objective). Greenhouse-Geisser estimates were used for all variables to account for deviations from sphericity. Post hoc t-tests with Bonferroni corrected p-values were completed for all significant effects. An alpha of 0.05 was applied for all other statistical tests.

Chapter 4: Results

Workload

Perceived Workload did not differ between Groups, $F(2, 48) = .30$, $p = .744$, $\eta_p^2 = .014$, but did vary between Blocks, $F(1.89, 77.6) = 11.35$, $p < .001$, $\eta_p^2 = .217$. Workload was steady between block one (M = 34.1, SEM = .806) and block two (M = 34.5, SEM = .809), t(43) = - .591, $p = .558$, Bonferroni corrected $\alpha = .017$), and decreased from block two to block three (M = 30.82, SEM = 1.02), $t(43) = 4.26$, $p < .001$. The interaction between Group and Block was not significant, $F(3.79, 77.6) = 1.79$, $p = .144$, $\eta_p^2 = .080$. The average score of the individual subscales for the NASA TLX workload measure are also provided in Figure 8 for visual reference.

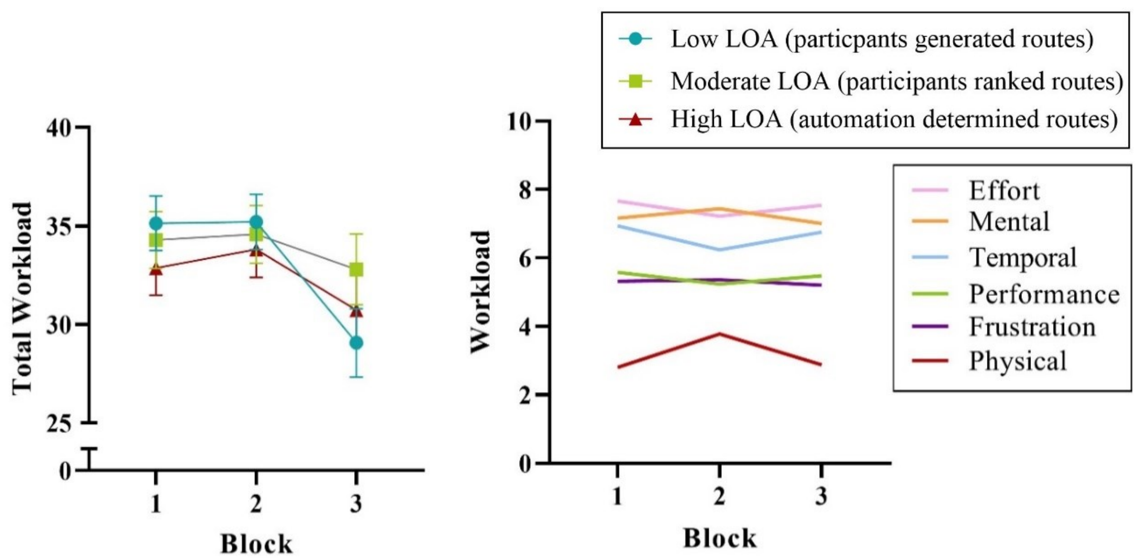


Figure 8. Perceived workload evaluated as summed raw scores on the six item, 10-point NASA task load index (min = 6, max = 60; left) and the average score of each subscale (right). The effort dimension asks the individual how hard they had to work to accomplish their level of performance. Mental demand asks how much mental activity was required (thinking, deciding, searching, calculating). Temporal demands refer to the rate of the task. Performance asks how successful the individual felt in accomplishing the task goals. Frustration level asks how insecure, or discouraged versus content or complacent the individual felt. Physical demands refer to activities including pushing, pulling, controlling. See Table 1 for detailed description. *Note: Error bars show the SEM.*

Situational Awareness

Situational awareness remained unchanged between Groups, $F(2, 41) = .43$, $p = .651$, $\eta_p^2 = .021$, and between Blocks, $F(1.97, 80.9) = .96$, $p = .385$, $\eta_p^2 = .023$. There was no interaction between Group and Block, $F(3.94, 80.9) = 1.35$, $p = .259$, $\eta_p^2 = .062$. In a post-hoc exploratory analysis, Person's correlation was used to determine whether SA was related to performance or visual sampling of critical cues at the individual level (Table 2). No significant correlations were noted.

Table 2. Pearson correlation coefficient between individual situational awareness and monitoring task performance, search task performance, and proportion of critical intervals (CIs) sampled across three experimental blocks.

Measure		Monitoring Performance			Search Performance			% CIs Fixated		
	Block	1	2	3	1	2	3	1	2	3
Situational	1	.154		.	.277	.	.	.085	.	.
Awareness	2	.	.159	.	.	.033	.	.	.045	.
	3	.	.	.191	.	.	.102	.	.	.212

Note: all ps > .05

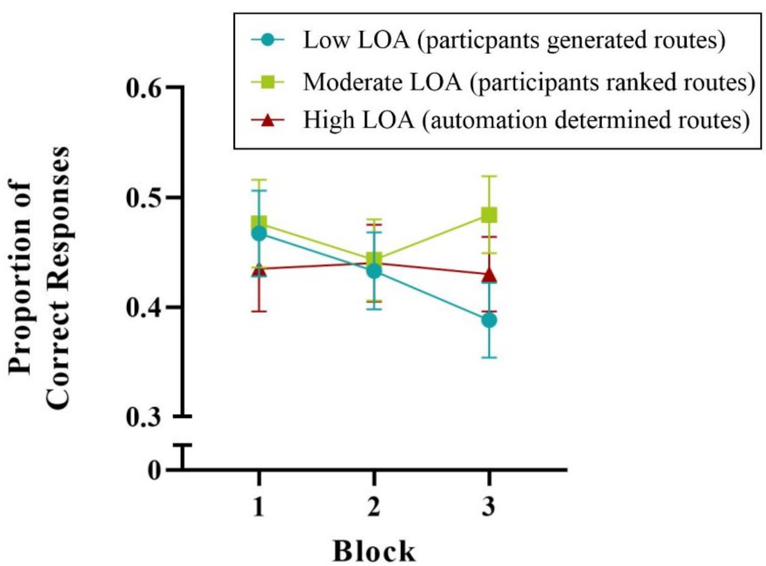


Figure 9. Proportion of correct responses examining situational awareness as evaluated with the situational awareness global assessment technique. *Note: Error bars show the SEM.*

Performance

Search Performance

Search performance differed between Groups, $F(2, 41) = 3.42$, $p = .042$, $\eta_p^2 = .143$. The performance was equal between the low LOA Group (M =.778, SEM = .022) and the moderate LOA Group (M =.790, SEM = .023), $t(27) = -.383$, $p = .705$), as well as between the moderate and high LOA (M =.843, SEM = .011) Groups, $t(27) = -2.156$, $p = .040$. The high LOA Group performed better than the low LOA group, $t(28) = -2.685$, $p = .012$, Bonferroni corrected $\alpha = .017$. However, differences between these groups was significant at baseline $t(28) = -2.677$, $p = .012$ and were therefor not produced by the experimental manipulation. Performance in the search task also differed between Blocks, $F(1.8, 73.9) = 16.46$, $p < .001$, $\eta_p^2 = .287$. Search performance increased from block one (M = .773, SEM = .015) to block two (M = .803, SEM = .012), t(43) = -3.207, p = .003, and from block two to block three (M = .832, SEM = .011), $t(43) = -3.130$, $p = .003$. There was no interaction between Group and Block, $F(3.61, 73.9) = 1.33$, $p = .268$, $\eta_p^2 = .061$.

Monitoring Performance

There was no difference in monitoring performance between Groups, $F(2, 41) = .29$, $p = .747$, $\eta_p^2 = .014$. There was a difference between Blocks, $F(1.93, 79) = 3.60$, $p = .034$, $\eta_p^2 = .081$. Performance decreased slightly, though not significantly, between blocks one (M = .952, SEM = .005) and two (M = .943, SEM = .004), t(43) = 1.96, p = .057, and increased from block two to three (M = .956, SEM = .004), $t(43) = -2.82$, $p =$

.007. There was no interaction between Group and Block, $F(3.86, 79.0) = .68$, $p = .610$, $\eta_p^2 = .032$.

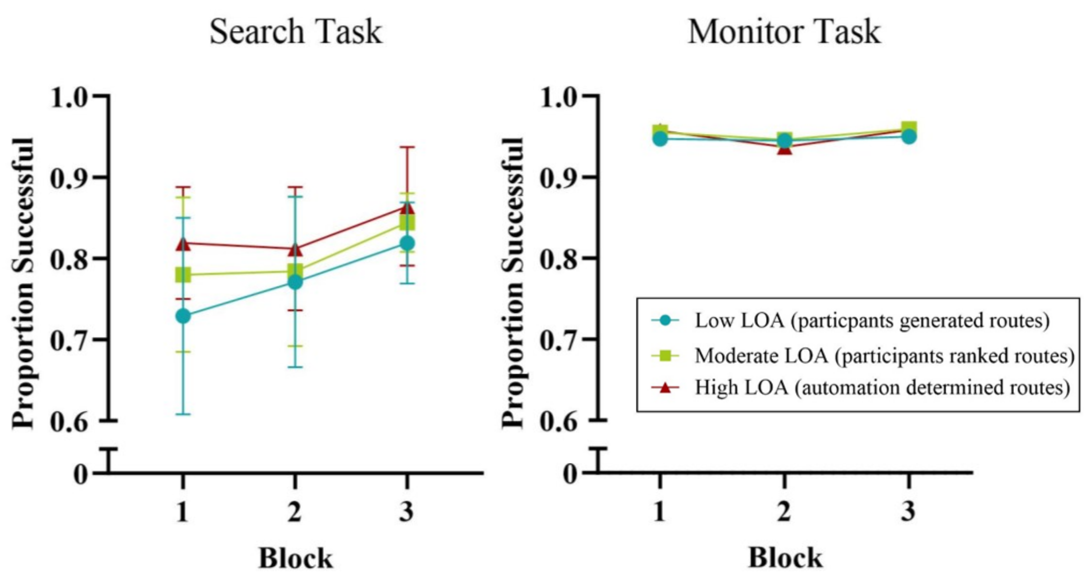


Figure 10. Proportion of correctly located targets (left) and segments successfully explored by the UV (right). *Note: Error bars show the SEM.*

Fixations

Monitoring ROIs

The proportion of dwell time in the monitoring ROIs were consistent between Groups, $F(2, 41) = .13$, $p = .879$, $\eta_p^2 = .006$, but differed across the Blocks, $F(1.82, 74.8) = 4.62$, $p = .015$, $\eta_p^2 = .101$, Dwell time decreased from block one (M = .272, SEM = .011) to block two (M = .240, SEM = .012), $t(43) = 3.29$, $p = .002$, and remained steady from block two to three (M = .250, SEM = .012), $t(43) = -.978$, $p = .334$. There was no interaction between Group and Block, $F(3.65, 74.8) = .84$, $p = .505$, $\eta_p^2 = .039$.

Search ROI

The proportion of dwell time in the search ROI was consistent between Groups, $F(2, 41) = .71$, $p = .497$, $\eta_p^2 = .034$. There was a main effect of Block, $F(1.92, 78.6) =$

3.37, p = .041, η_p^2 = .076 but differences were not supported in post hoc evaluations (B1 M = .580, SEM = .013; B2 M = .546, SEM = .020; B3 M = .573, SEM = .015; ps > .024). There was no interaction between Group and Block, $F(3.83, 78.6)$ = .66, p = .614, η_p^2 = .031.

Critical Interval

The proportion of ROIs fixated during critical intervals, where the UVs were within their two-unit range to clear risky areas, was consistent between Groups, $F(2, 41)$ = .163, $p = .850$, η_p^2 = .008. Fixated ROIs in these intervals difference across the Blocks, $F(1.94, 79.37)$ = 7.43, $p < .001$, η_p^2 = .153, where they decreased from block one (M = .855, SEM = .022) to block two (M = .760, SEM = .031), $t(43)$ = 3.96, p < .001 and then remained steady from block two to block three (M = .809, SEM = .027), $t(43)$ = -2.05, p = .047. There was no interaction between Group and Block, $F(3.87, 79.4)$ = .98, p = .424, η_p^2 = .046.

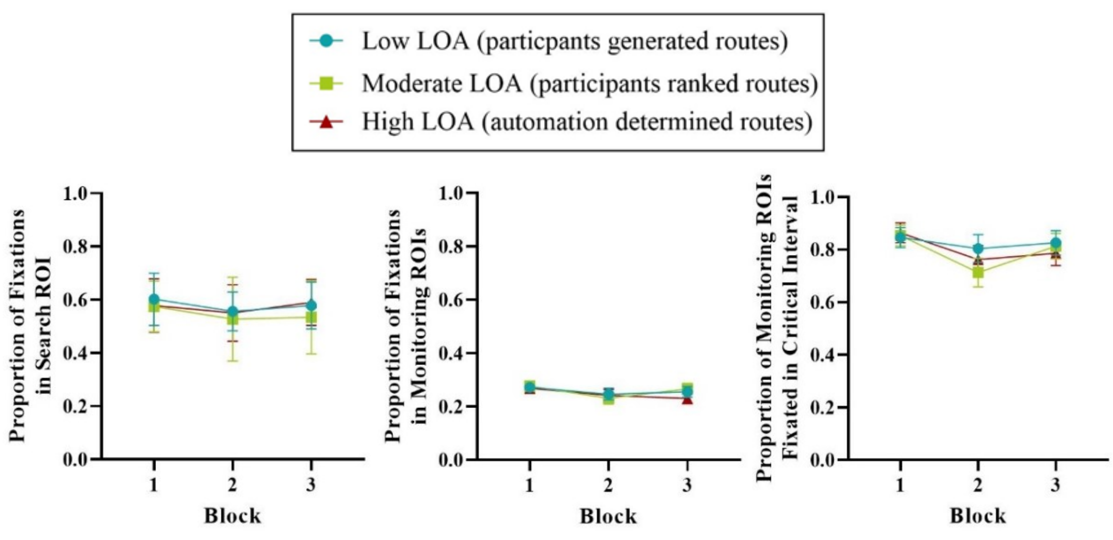


Figure 11. Proportion of fixation dwell time in search (left) and monitoring (middle) ROIs, and the proportion of fixations in monitoring ROIs that were fixated when the UVs were approaching (within two units) a risky area (right). *Note 1: ROI(s) = Region(s) of Interest. Note 2: Error bars show the SEM.*

Saccades

Between Search ROI and Monitoring ROIs

The proportion of saccades between the task regions was equal across Groups, $F(2, 41) = .300$, $p = .743$, $\eta_p^2 = .014$. There was a change across the Blocks, $F(1.89, 77.56) = 19.61$, $p < .001$, $\eta_p^2 = .323$, where the proportion of saccades between tasks increased from block one (M = .659, SEM = .011) to block two (M = .699, SEM = .012), $t(43) = 4.96$, $p < .001$, and from block two to block three (M = .716, SEM = .012), $t(43) = 1.827$, $p = .037$. There was no interaction between Group and Block, $F(3.78, 77.56) = .32$, $p = .857$, $\eta_p^2 = .015$. Saccades between the Monitoring ROIs were inverse to the saccades between the search and monitoring ROIs and therefore decreased across each experimental block (Figure 12).

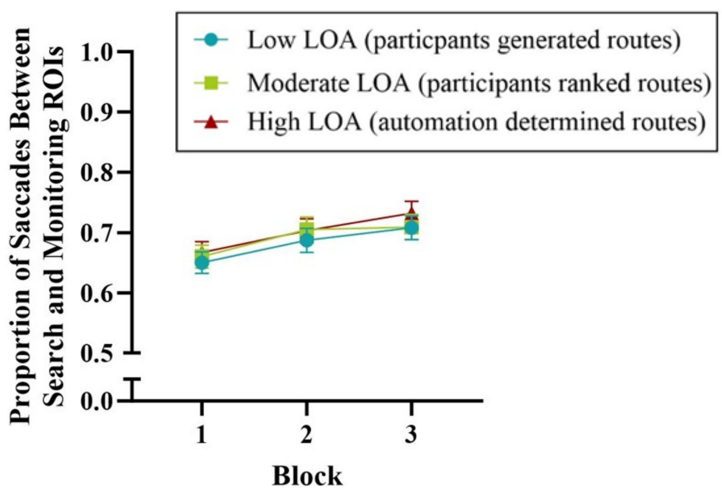


Figure 12. Proportion of saccades that occurred between monitoring ROIs and the search ROI. *Note 1: ROI = Region(s) of Interest. Note 2: Error bars show the SEM.*

Chapter 5: Discussion

The purpose of this study was to evaluate whether involvement in the process of planning a UVs route affects a human operator's ability to monitor the same UV. To examine this relationship, the primary objective of this study was to do determine whether SA, situation assessment, perceived workload, and performance varied when a human completed the route generation (low LOA group) or the route selection (moderate LOA) functions, or neither (high LOA). When participants were assisted by a planning system that operated at a low or moderate LOA, perceived workload was expected to increase, while performance, SA and visual sampling in critical intervals were expected to improve compared to trials where they were assisted by a system operating at a high LOA. To evaluate whether periods of involvement in the planning process, where the system operates at a low or moderate LOA, results in lasting changes to the humans monitoring abilities (e.g., mental model or visual sampling strategy), the secondary objective was to examine the same variables when the system returned to a high LOA.

Between block one and block two there was a significant decrease in the proportion of dwell time in the monitoring ROIs, an increased in the saccades between tasks indicating less saccades between the monitoring ROIs, as well as a decrease in the proportion of critical intervals sampled for all groups, demonstrating a shift in visual sampling away from the monitoring task regardless of the automation level in the planning phase. The shift in visual sampling appeared to be beneficial as performance in the search task improved in block two. When the planning system returned to a high LOA in block three) the proportion of saccades that were between the task ROIs continued to increase, perceived workload decreased, and performance in both the search and monitoring tasks improved. Together these results indicate that although participants allocated less visual attention to the monitoring task throughout the session, their strategy was successful as they maintained a high level of performance in monitoring and improved in the search task. Critically however, these effects were not modified by group indicating that manipulating the automation level of the route planning system did not affect performance, perceived workload, situational awareness, or situational assessment. With no between group differences in block two, the hypotheses were not supported, and the secondary objective could not be explored.

Workload and LOA

The automation levels of the route planning system evaluated here aimed to assess whether the specific processes of route generation and route selection are fundamental to the development of situational awareness. When the system operated at a low or moderate LOA participants in these groups were expected to allocate more cognitive resources to comparing potential routes and report a higher perceived workload. Instead, perceived workload remained steady from block one to block two for all groups. This was surprising as participants in the low and moderate LOA groups had an objectively higher workload. Since perceived workload was evaluated as a single value for the entire block, the increased workload in planning may not have been proportionality large enough to influence the overall workload scores. Differences may have emerged in the planning phase, but this could not be extracted from the current dataset.

When participants in the low and moderate LOA groups returned to the highly automated planning system in block three, their workload was expected to decrease to initial levels. Further, the anticipated inverse relationship between the LOA of the planning system and perceived workload was expected to be driven by the workload dimensions linked to increased cognitive processing - effort and mental demands. Instead, overall workload was perceived to decrease for all groups, and appears to have been driven by physical demand (Figure 8). This was unexpected as route planning was not intended, as, nor perceived to be, a physically demanding task at any LOA – in their rating of the workload dimensions participants rated physical demand the lowest.

While it is unclear why physical demand was perceived to change, first it can be considered why cognitive demands did not change. While planning with the system at a low LOA participants were instructed to consider and minimize both the overall route distance and distance traveled in risky areas. While these parameters would make the UVs route more efficient, they did not provide any indication as to whether the UV's path would be clear and were not directly related to the participants role in the monitoring task. Therefore, there was no incentive for participants to put effort into performing these tasks well. With minimal integration of information necessary and without a standard for route quality, it appears that the route generation and selection functions were not viewed as effortful or mentally demanding compared to the remainder of the task. One parameter that could have linked the planning and monitoring stages was the relative timing of CIs in segments that were monitored concurrently. Unfortunately, this was not accounted for in the study design and segments were planned individually. While it is possible that some participants considered the relative timing of the CIs in planning the second segment, this was not verbalized to the researcher as a consideration by any participant and it would require the memorization of the route of the first segment for each trial – a challenging feat.

Although the change in perceived workload only reached statistical significance between blocks two and three, the perceived physical workload appears to have followed the anticipated trend by increasing from block one to block two and decreasing from block two to block three. The perceived increase in physical demand in the second experimental block may be attributed to transitions to the adjacent monitor for the planning exercises. The rotation required was less than 45° (~ 30°) and could be achieved either by rotating one's chair toward the monitor, rotating one's head, or a combination of each. While a single rotation requires minimal physical effort for an able-bodied adult, it is possible that the forty-four (there and back for each trial) additional rotations were perceived as an increase in physical demands when evaluated relative to blocks one and three. Simply wearing the eye-tracker for the additional time to complete the planning exercises, rather than the planning exercises themselves, may have also led to the increase in perceived physical demand. While the EyeLink®II head-mounted system is not heavy alone, the additional weight of the gear may have placed strain on the neck muscles that was perceived as a change in physical demand. While there are no studies on the physical impacts of wearing a head mounted eye-tracker know to the author, one study on the physiological and psychological neck loads of ballistic helmets did examine the impacts of night vision devices (Kim & Jeong, 2020). Like the EyeLink®II, the weight of night vision devices is distributed anteriorly. The addition of night vision devices was noted to increase both subjective neck load and neck pain (Kim & Jeong, 2020). The physical exertion, or discomfort associated with anteriorly distributed weight of the eye-tracker in this study may have been expressed as physical workload by the participants.

The remaining physical tasks required in the simulation, controlling a mouse and minimal head movements to look at the screen, are unlikely to contribute to perceived physical demand. Alternatively, the decrease in overall workload might simply reflect the relatively shorted length of the third block which may have been categorized as physical workload simply by default. Regardless of the mechanism, with no difference between LOA groups in the perceived cognitive demands between levels of automation, the planning manipulation does not appear to have elicited additional assessment of the segments. Consequently, since the level of automation did not result in additional assessment of the segments, or increase the mental workload in the planning phase, it is not surprising that no effects emerged between groups for the remaining variables of performance, SA, or visual sampling. Reported findings, particularly those regarding the impacts of a planning systems LOA, should be interpreted cautiously.

Future examinations of the impacts of automation level and workload on situational awareness across task phases should ensure that information relevant to the monitoring phase is considered in the planning phase. The impact of a planning systems automation level on specific dimensions of workload should also be considered. While the relevance of each workload dimension is expected to vary across settings, future investigations may benefit from using the weighted NASA-TLX protocol in pilot testing to confirm that the intended dimension of workload is targeted in manipulations of a systems LOA.

Situational Awareness

On average, participants answered 44% (SE = 1.8%) of the situational awareness queries correctly. This moderate performance was stable across blocks and between groups. The lack of association between individual situational awareness and both performance (Endsley, 1995a) and visual sampling of critical stimuli (Eisma et al., 2018; Table 2) suggests that either there is no relationship between SA and performance or situation assessment, or that the methods applied here did not capture the relationships. With a strong theoretical link between SA and performance, as well as SA and situation assessment, it is plausible that the SA assessment technique applied here was not appropriate. While it is recommended to conduct a goal-based task analysis, with input from subject matter experts, to identify SAGAT queries to be utilized (Endsley & Garland, 2000), this was impractical for the low-fidelity simulation employed here. Instead, questions were selected simply to reflect the three levels of situational awareness. Without consideration of task goals in corresponding applied settings, some questions may have been inappropriate to assess SA. For example, participants were queried on whether the UV had to make a turn before the end of the trial. This knowledge was not relevant to their goal to locate targets and ensure the UVs did not enter the risky areas, and an incorrect answer does not indicate that they are unable to meet their task goals.

The SAGAT has been employed successfully in health care (Dishman et al., 2020; Lavoie et al., 2016), military (Strater et al., 2006) and aerospace (Endsley & Garland, 2000; O'Brien & O'Hare, 2007; Zhou et al., 2022) domains. While the SAGAT does not necessitate high fidelity simulations (Endsley & Garland, 2000), in these domains queries could be readily linked to the goal of the tasks. For example, queries developed to assess SA of nursing students have included whether there is a leak around the patient's face mask supplying oxygen, what the patient's breath sounds indicate, and whether the patient requires additional respiratory support or intervention (Dishman et al., 2020). These questions both capture the levels of SA and are based in information that is explicitly necessary to providing adequate patient care. To properly assess situational awareness of UV route planning and subsequent monitoring, it may be beneficial to conduct a thorough goal-based task analysis and validation of SAGAT adaptations, such as the ones completed in the healthcare domains (Dishman et al., 2020; Lavoie et al., 2016) prior to further research.

Visual Sampling and Performance

Over the course of the three blocks, a subtle but statistically significant shift in visual sampling strategy emerged. A decrease in the proportion of dwell time spent in monitoring ROIs indicates that less visual attention was allocated to monitoring the UVs as the blocks progressed. Although less dwell time in the monitoring task would be expected to result in more dwell time in the search ROI, this was not observed. The only on-screen cues outside the analyzed ROIs were the “CANCEL” buttons displayed under each monitoring route, and the timer for the search task. Since the cancel buttons do not change or offer information relevant to the task, it is likely that the proportion of dwell time lost from the monitoring task was allocated to the search task through increased sampling of the timer.

Sampling strategies also shifted to prioritize more frequent sampling of the search ROI. Saccades between the search and monitoring ROIs increased across each block – upon completing a fixation in a monitoring ROI, participants were more likely to direct attention to the search task than the other monitoring ROI. The strategy was effective – performance in both the monitoring and search tasks either remained equal or improved.

The dwell time and saccades align with previous research that has noted participants learning to allocate visual attention, in proportion of overall dwell time, more appropriately within a study session (Eisma et al., 2018). With more dwell time and more frequent assessments of the search ROI, it is logical that performance in the search improved across the blocks. However, it was surprising that performance in the monitoring task improved while the proportion of fixated CIs decreased.

Within the monitoring task, cancellation of a UV's route was only necessary for one in every eight segments (12.5%). Without any action taken participants were expected to successfully complete the remaining 87.5% of the trials. Each correct cancelation of a UV would increase this performance by 2.5%. An average performance in monitoring ranging from 94.3% to 95.6% across the blocks, indicate that participants likely canceled three out of five UVs correctly per block throughout the study – though the proportion of correct trials may have also been achieved with a greater number of UV cancellations coupled with incorrectly cancelling a UV that was not approaching a risky area. Nevertheless, though a statistically significant difference was noted between blocks two and three, the practical significance of this 1.3% improvement amounted to less than one additional successful segment. Since there were no gains in SA and the visual sampling of this task decreased, it is unlikely that the minimal improvement in monitoring task can be attributed to changes in situational awareness. Instead, this improvement may reflect other task relevant learning, such as improvements in procedural knowledge.

It should be noted that the segment and UV stimuli were created in MATLAB and imported into Experiment Builder as video files, where it was displayed during data collection. This approach introduced the potential for error in the timing accuracy of the visual data collected relative to the stimuli presented through the simulation. To evaluate whether a CI was fixated, data from the two systems was matched with the video frames. The CI was identified by the video frames where the UV was within the two grid squares of each risky area when the videos were created in MATLAB. Throughout data collection, visual samples were stored relative to the simulation frame by Experiment Builder. In data processing, the simulation frame of each fixation was compared to the CI frames extracted from MATLAB to determine whether a fixation occurred in that ROI during the CI. Timing inconsistencies between the frame that graphics were rendered on screen in video creation and the presentation of the videos during the simulation could have led to fixation data that was improperly matched to the corresponding frames of the CI. While it would be preferred to avoid this source of error, the use of both MATLAB and Experiment Builder to develop the simulation arose from the limitations of each program which are discussed further below.

The impact of eye tracking hardware on participants within the study procedure should also be considered. For this study, participants were asked to wear the head mounted Eyelink®II for approximately ninety minutes across three blocks. While the head-mounted set up is described as comfortable (SR Research, 2005), many participants expressed discomfort. While all efforts were made to reduce discomfort, including re-positioning the straps and additional padding, participants typically expressed feeling pressure on their forehead or temples during the second or third experimental block. To mitigate these sensations, participant-paced breaks between blocks often extended beyond the anticipated five minutes. Though the discomfort was described as mild, and all participants were able to wear the eye tracker for the entirety of the study protocol, future studies should consider minimizing the time required to wear a head-mounted eye tracker or employ a desk mounted configuration when longer continuous measurements are necessary. In future applications of the head-mounted hardware, consideration should also be given to the timing of tasks that require, or do not require, evaluating visual sampling. For example, in the present study the eye-tracker was worn for the SA assessment even though visual data was not collected for this period. To reduce discomfort, desk-mounted eye tracking could have been used, where participants would have been able to adjust their head position during the SAGAT queries and the planning phase of each trial. In the future, studies should give considerable thought to the experience of the participants that must wear a head-mounted configuration and use time wearing the eye-tracker efficiently.

Overall, the findings of this study suggest that in basic UV operations the automation level of a route planning system to determine the UVs route does not impact one's ability to subsequently monitor the UV. Likewise, the uniformity of situational awareness and situation assessment measures across experimental groups suggest that engaging in the route generation and selection functions of the planning phase will not improve an operator's ability to gain information from the simulation environment. With no decrements from the route planning system operating at a high LOA, it may be preferable to highly automate the route planning of UVs to increase efficiency while reducing the tasks required of a human operator. However, given the limitations of this study, these interpretations should be taken with a high level of caution. Importantly, these findings are limited to monitoring the position of a virtual UV in low-fidelity environments where information available in the planning phase provides minimal insight on changes that may occur in the monitoring phase.

Limitations and Future Directions

In addition to the shortfalls of the LOA manipulation, and the SA assessment technique, limitations of the study design employed include software compatibility, and potential incongruencies between the task demands of the simulation and the true demands in UV monitoring settings. Each are discussed with recommendations for future studies.

Upon conception, the stimulation was intended to be programmed in MATLAB with the Psychtoolbox extension. The extension provides a packaged set of functions built for stimulus presentation and response collection in MATLAB and Octave (Psychtoolbox, n.d.). In the development of the simulation in MATLAB, a timing error emerged between the Display PC and Host PC that prevented further processing of the script and could not be rectified. While Psychtoolbox is listed by SR Research as a compatible stimulus presentation software, MATLAB itself is not included as a compatible programming language (SR Research, n.d.-a). Without Psychtoolbox, MATLAB could not be interfaced with the EyeLink®II. Simulation programming was then transitioned to Experiment Builder software developed by SR Research.

Experiment Builder for programming and Data Viewer for data processing, are highly compatible with the EyeLink®II. While Data Viewer was sufficient for the data preparation and visualization required in this project, the drag and drop Experiment Builder interface was ill-equipped for the development of this simulation. For example, in the development of the simulation lapses in rendering the simple UV videos were noted.

Videos of the UVs for the monitoring task would freeze momentarily and then resume numerous frames ahead of the last rendered frames. This problem was remedied by installing additional RAM (random access memory) to a total of 64 GB (gigabyte; 4x16 GB cards), a requirement that far exceeds the indicated four GBs recommended for Experiment Builder by SR Research (SR Research, n.d.-b). This high RAM requirement to display the simple videos suggests inefficient computing in the Experiment Builder program. The resource demands are particularly surprising since all background activities are halted on the Display PC while Experiment Builder projects are running. Because the basic stimulus presentation approached the processing limits of Experiment Builder, a third computer was employed for route planning tasks, workload assessments, and SA queries. To pair trials between computers, the Display PC indicated a trial ID that was manually entered by the researcher into the adjacent computer. This process resulted in inconsistent delays between the occlusion of stimuli and the presentation of the SAGAT queries, potentially introducing error in the measure of SA; increased the duration of the overall study; and increased the duration that the eye tracker was worn, contributing to participant discomfort.

In addition to inefficient computations, limitations of Experiment Builder include the lack of capability to execute sections of a project in isolation – meaning the entire project had to run to confirm any changes – and a tedious process to adjust parameters. While recommending a particular programming language is outside the scope of this study, it is recommended that future studies collaborate with a subject matter expert or the SR Research support team to determine the best software to meet study requirements.

While the simulation employed in this study was anticipated to mimic the basic demands of evaluating information provided by a UV while monitoring the UVs position, the developed program fell short of this goal. It is unlikely that the minimal features offered for consideration in the monitoring task captured the relationship between planning and monitoring in an applied setting, nor the goals of either tasks. To approximate these task goals in future studies it could be helpful to conduct goal-based task analyses across numerous UV operation settings, identify themes, and develop a simulation that captures recurrent goals. Since varying the automation level of a system can be unique at each function (i.e., monitoring, generating, selecting, and implementing; Kaber & Endsley, 1997) identifying the goals and subgoals for each of these functions would inform the which LOAs should be examined further and would allow for the appropriate adaption of SAGAT queries to match task goals. A goal-based analysis is preferred to a general task-oriented analysis to examine the process of integrating information to reach task goals across diverse settings (Endsley & Garland, 2000). Such an analysis may also shine light on the typical format of information and cues that must be perceived in UV route planning, which can be used to inform apparatus design and assessment of visual sampling in future research.

While automated planning systems are employed to assist humans in the deployment of UVs, the impacts on situational awareness and overall performance remain unknown. Ongoing studies are required to determine the optimal automation level for planning systems across applied domains. Limitations of this study can be used to inform the development of future investigations. Of particular importance is the establishment of task goals in settings where UVs are deployed, which can be captured in goal-based task analysis. SA, visual sampling, and performance should then be evaluated in simulations that capture these goals to understand the extent, and mechanisms, in which automated planning systems can impact human teammates.

Chapter 6: Conclusion

Uncrewed vehicles can extend operational limits (Balaram et al., 2021; de Alcantara Andrade et al., 2019; Haddal & Gertler, 2010) and perform well under normal operations. However, unanticipated situations, where the appropriate response is beyond the UVs programmed capability, can lead to – sometimes catastrophic – failures. To mitigate the risks of unanticipated situations, human operators are often tasked to monitor the UVs and intervene as necessary in situations the UV is not equipped to handle. Automated route planning systems can be used in conjunction with UVs to reduce the workload of operators, allowing them to focus on monitoring. However, removing the operator from the planning process may limit their SA of the UV and the operational environment to be monitored. Since current SA is the basis for building and maintaining SA, the development of SA in the planning process is of particular interest. The use of automated planned systems may affect an operator's ability to initially build and then maintain SA, which in turn may impact their ability to monitor the UV.

This study aimed to evaluated whether an operator's involvement in the route planning process contributes to their ability to monitor UVs. To examine this, a non-descript simulation of a UV operation was developed. Planning functions were allocated to either a human operator or an automated system, as denoted by the systems level of automation. Situational awareness, situation assessment, perceived workload, and performance were evaluated when the operator was monitoring the UV. Results suggest that the automation level of a planning system does not impact any of the variables examined, and therefore does not impact one's ability to monitor UVs. However, this conclusion should be interpreted cautiously given the limitations of the study. With no difference between experimental groups, the secondary objective – to evaluated whether benefits of an operator engaging in planning would remain if planning was later allocated entirely to the system – could not be evaluated. To overcome the limitations of the present study, future investigations should be conducted with simulations that reflect the specific task goals of UV operations, which can be captured in goal-based task analyses. The goal-based tasks should also be used to inform the development of questions to appropriately assess SA. Finally, subject matter experts should be consulted to ensure the technology employed can efficiently meet study requirements. With these changes incorporated SA, situation assessment, perceived workload, and performance should then be re-evaluated to understand the extent, and mechanisms, in which an automated route planning systems might impact a human operator's ability to monitor UVs. Results of this, and further investigations, may be used to inform the allocation of tasks to a human or system to optimize performance in UV operations.

References

Alawad, W., Halima, N. B., & Aziz, L. (2023). An Unmanned Aerial Vehicle (UAV) System for Disaster and Crisis Management in Smart Cities. *Electronics*, *12*(4), Article 4. https://doi.org/10.3390/electronics12041051

Al-Kaff, A., Gómez-Silva, M. J., Moreno, F. M., de la Escalera, A., & Armingol, J. M. (2019). An Appearance-Based Tracking Algorithm for Aerial Search and Rescue Purposes. *Sensors*, *19*(3), Article 3. https://doi.org/10.3390/s19030652

Bai, X., Li, B., Xu, X., & Xiao, Y. (2022). A Review of Current Research and Advances in Unmanned Surface Vehicles. *Journal of Marine Science and Application*, *21*(2), 47–58. https://doi.org/10.1007/s11804-022-00276-9

Balaram, J., Aung, M., & Golombek, M. P. (2021). The Ingenuity Helicopter on the Perseverance Rover. *Space Science Reviews*, *217*(4), 56. https://doi.org/10.1007/s11214-021-00815-w

Balfe, N., Sharples, S., & Wilson, J. R. (2018). Understanding Is Key: An Analysis of Factors Pertaining to Trust in a Real-World Automation System. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *60*(4), 477–495. https://doi.org/10.1177/0018720818761256

Banbury, S., Forbes, K., & Callian Ltd. (2022). *Policy Considerations for Future Autonomous Systems*. Defence Research and Development Canada. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://cradpdf.drdc-rddc.gc.ca/PDFS/unc389/p814837_A1b.pdf

Banbury, S., & Tremblay, S. (2004). SPAM: The Real-Time Assessment of SA. In *A Cognitive Approach to Situation Awareness: Theory and Application*. Routledge.

Blasch, E., Kadar, I., Salerno, J., Kokar, M. M., Das, S., Powell, G. M., Corkill, D. D., & Ruspini, E. H. (2006). Issues and challenges of knowledge representation and reasoning methods in situation assessment (Level 2 Fusion). *Signal Processing, Sensor Fusion, and Target Recognition XV*, *6235*, 355–368. https://doi.org/10.1117/12.669779

Blum, M., & Vempala, S. (2020). The complexity of human computation via a concrete model with an application to passwords. *Proceedings of the National Academy of Sciences*, *117*(17), 9208–9215. https://doi.org/10.1073/pnas.1801839117

Bouchacourt, F., & Buschman, T. J. (2019). A Flexible Model of Working Memory. *Neuron*, *103*(1), 147-160.e8. https://doi.org/10.1016/j.neuron.2019.04.020

Busby, S., & Witucki-Brown, J. (2011). Theory Development for Situational Awareness in Multi-casualty Incidents. *Journal of Emergency Nursing*, *37*(5), 444–452. https://doi.org/10.1016/j.jen.2010.07.023

Bustamante, E., & Spain, R. (2008). Measurement Invariance of the Nasa TLX. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52*, 1522–1526. https://doi.org/10.1177/154193120805201946

Chauvin, C., Clostermann, J.-P., & Hoc, J.-M. (2008). Situation Awareness and the Decision-Making Process in a Dynamic Situation: Avoiding Collisions at Sea. *Journal of Cognitive Engineering and Decision Making*, *2*, 1–23. https://doi.org/10.1518/155534308X284345

Chen, X., Tang, J., & Lao, S. (2020). Review of Unmanned Aerial Vehicle Swarm Communication Architectures and Routing Protocols. *Applied Sciences*, *10*(10), Article 10. https://doi.org/10.3390/app10103661

Clark, H., McLaughlin, A. C., & Feng, J. (2017). Situational Awareness and Time to Takeover: Exploring an Alternative Method to Measure Engagement with High-Level Automation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1452–1456. https://doi.org/10.1177/1541931213601848

Cognolato, M., Atzori, M., & Müller, H. (2018). Head-mounted eye gaze tracking devices: An overview of modern devices and recent advances. *Journal of Rehabilitation and Assistive Technologies Engineering*, *5*, 205566831877399. https://doi.org/10.1177/2055668318773991

Colebank, J. L. (2008). *Automated Intelligent Agents: Are They Trusted Members of Military Teams?* NAVAL POSTGRADUATE SCHOOL MONTEREY CA.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Van Essen, D. C., & Shulman, G. L. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761–773. https://doi.org/10.1016/s0896-6273(00)80593-0

Cummings, M. L., Buchin, M., Carrigan, G., & Donmez, B. (2010). Supporting intelligent and trustworthy maritime path planning decisions. *International Journal of Human-Computer Studies*, *68*(10), 616–626. https://doi.org/10.1016/j.ijhcs.2010.05.002

Cummings, M. L., Marquez, J. J., & Roy, N. (2012). Human-automated path planning optimization and decision support. *International Journal of Human-Computer Studies*, *70*(2), 116–128. https://doi.org/10.1016/j.ijhcs.2011.10.001

de Alcantara Andrade, F. A., Reinier Hovenburg, A., Netto de Lima, L., Dahlin Rodin, C., Johansen, T. A., Storvold, R., Moraes Correia, C. A., & Barreto Haddad, D. (2019). Autonomous Unmanned Aerial Vehicles in Search and Rescue Missions Using Real-Time Cooperative Model Predictive Control. *Sensors (Basel, Switzerland)*, *19*(19), 4067. https://doi.org/10.3390/s19194067

De Winter, J. C. F., Eisma, Y. B., Cabrall, C. D. D., Hancock, P. A., & Stanton, N. A. (2019). Situation awareness based on eye movements in relation to the task environment. *Cognition, Technology & Work*, *21*(1), 99–111. https://doi.org/10.1007/s10111-018-0527-6

De Wit, P. A. J. M., & Moraes Cruz, R. (2019). Learning from AF447: Human-machine interaction. *Safety Science*, *112*, 48–56. https://doi.org/10.1016/j.ssci.2018.10.009

Di Flumeri, G., De Crescenzio, F., Berberian, B., Ohneiser, O., Kramer, J., Aricò, P., Borghini, G., Babiloni, F., Bagassi, S., & Piastra, S. (2019). Brain–Computer Interface-Based Adaptive Automation to Prevent Out-Of-The-Loop Phenomenon in Air Traffic Controllers Dealing With Highly Automated Systems. *Frontiers in Human Neuroscience*, *13*. https://www.frontiersin.org/articles/10.3389/fnhum.2019.00296

Dishman, D., Fallacaro, M. D., Damico, N., & Wright, M. C. (2020). Adaptation and Validation of the Situation Awareness Global Assessment Technique for Nurse Anesthesia Graduate Students. *Clinical Simulation in Nursing*, *43*, 35–43. https://doi.org/10.1016/j.ecns.2020.02.003

Duchowski, A. (2007). *Eye Tracking Methodology*. Springer. https://doi.org/10.1007/978-1-84628-609-4

Durso, F. T., Hackworth, C. A., & Truitt, T. R. (1999). *Situation awareness as a predictor of performance in en route air traffic controllers.: (402702004-001)* [Data set]. https://doi.org/10.1037/e402702004-001

Eisma, Y. B., Cabrall, C. D. D., & De Winter, J. C. F. (2018). Visual Sampling Processes Revisited: Replicating and Extending Senders (1983) Using Modern Eye-Tracking Equipment. *IEEE Transactions on Human-Machine Systems*, *48*(5), 526–540. https://doi.org/10.1109/THMS.2018.2806200

Endsley, M. R. (1993). A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters. *The International Journal of Aviation Psychology*, *3*(2), 157–168. https://doi.org/10.1207/s15327108ijap0302_5

Endsley, M. R. (1995a). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human Factors*, *37*, 65. https://doi.org/10.1518/001872095779049499

Endsley, M. R. (2011). Direct Measurement of Situation Awareness: Validity and Use of SAGAT. In E. Salas & A. S. Dietz (Eds.), *Situational Awareness*. Routledge.

Endsley, M. R. (2015). Situation Awareness Misconceptions and Misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4–32. https://doi.org/10.1177/1555343415572631

Endsley, M., R. (2019). *Human Factors & Aviation Safety: Testimony to the United States House of Representatives Hearing on Boeing 737-Max8 Crashes*.

Endsley, M. R. (2020). The Divergence of Objective and Subjective Situation Awareness: A Meta-Analysis. *Journal of Cognitive Engineering and Decision Making*. https://doi.org/10.1177/1555343419874248

Endsley, M. R. (2021). A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. *Human Factors*, *63*(1), 124–150. https://doi.org/10.1177/0018720819875376

Endsley, M. R., & Garland, D. J. (2000). *Situation awareness analysis and measurement* (pp. xxiii, 383). Lawrence Erlbaum Associates Publishers.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–492. https://doi.org/10.1080/001401399185595

Endsley, M. R., & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381–394. https://doi.org/10.1518/001872095779064555

Esmaeili, B. (2017). Measuring the Impacts of Safety Knowledge on Construction Workers' Attentional Allocation and Hazard Detection Using Remote Eye-Tracking Technology. *Journal of Management in Engineering*, *33*. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000526

Fitts, P. (1951). *HUMAN ENGINEERING FOR AN EFFECTIVE AIR-NAVIGATION AND TRAFFIC-CONTROL SYSTEM*. National Research Council. https://apps.dtic.mil/sti/citations/ADB815893

Gafurov, S. A., & Klochkov, E. V. (2015). Autonomous unmanned underwater vehicles development tendencies. *Procedia Engineering*, *106*, 141–148.

Haddal, C. C., & Gertler, J. (2010). *Homeland Security: Unmanned Aerial Vehicles and Border Surveillance*. https://apps.dtic.mil/sti/citations/ADA524297

Hart, S. G. (2006). *Nasa-Task Load Index (NASA-TLX); 20 Years Later*. HUMAN FACTORS AND ERGONOMICS SOCIETY 50th ANNUAL MEETING.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hasanzadeh, S., Esmaeili, B., & Dodd, M. (2017). Impact of Construction Workers' Hazard Identification Skills on Their Visual Attention. *Journal of Construction Engineering and Management*, *143*. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001373

Hasanzadeh, S., Esmaeili, B., & Dodd, M. D. (2018). Examining the Relationship between Construction Workers' Visual Attention and Situation Awareness under Fall and Tripping Hazard Conditions: Using Mobile Eye Tracking. *Journal of Construction Engineering and Management*, *144*(7), 04018060. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001516

Hilario, J., Penaloza, C., Hernandez-Carmona, D., Balbuena, J., Quiroz, D., Ramirez, J., & Cuellar, F. (2019). Development of a Mobile Robot for Inspection of Analog Gauges in Industrial Plants Using Computer Vision. *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, 209–214. https://doi.org/10.1109/CIS-RAM47153.2019.9095800

Ho, G., Pavlovic, N., Arrabito, R., & Abdalla, R. (2011). *Human Factors Issues When Operating Underwater Remotely Operated Vehicles and Autonomous Underwater Vehicles* [Technical Report]. Defence R&D Canada.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795. https://doi.org/10.3758/bf03206794

Hollister, D. L., Gonzalez, A., & Hollister, J. (2017). Contextual Reasoning in Human Cognition and the Implications for Artificial Intelligence Systems. In P. Brézillon, R. Turner, & C. Penco (Eds.), *Modeling and Using Context* (pp. 599–608). Springer International Publishing. https://doi.org/10.1007/978-3-319-57837-8_48

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied*, *12*, 67–78. https://doi.org/10.1037/1076-898X.12.2.67

Inagaki, T. (2003). Adaptive automation: Sharing and trading of control. In *Handbook of cognitive task design* (pp. 147–169). Lawrence Erlbaum Associates Publishers. https://doi.org/10.1201/9781410607775.ch8

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 307–317. https://doi.org/10.1037/0278-7393.18.2.307

Jensen, M. S., Yao, R., Street, W. N., & Simons, D. J. (2011). Change blindness and inattentional blindness. *WIREs Cognitive Science*, *2*(5), 529–546. https://doi.org/10.1002/wcs.130

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Kaber, D. B., & Endsley, M. R. (1997). Out-of-the-loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. *Process Safety Progress*, *16*(3), 126–131. https://doi.org/10.1002/prs.680160304

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, *5*, 113–153. https://doi.org/10.1080/1463922021000054335

Kim, S., & Jeong, W. (2020). Physiological and psychological neck load imposed by ballistic helmets during simulated military activities. *Fashion and Textiles*, *7*(1), 27. https://doi.org/10.1186/s40691-020-00216-7

Knocton, S., Hunter, A., Connors, W., Dithurbide, L., & Neyedli, H. F. (2023). The Effect of Informing Participants of the Response Bias of an Automated Target Recognition System on Trust and Reliance Behavior. *Human Factors*, *65*(2), 189–199. https://doi.org/10.1177/00187208211021711

Laghari, A. A., Jumani, A. K., Laghari, R. A., & Nawaz, H. (2023). Unmanned aerial vehicles: A review. *Cognitive Robotics*, *3*, 8–22. https://doi.org/10.1016/j.cogr.2022.12.004

Lavoie, P., Cossette, S., & Pepin, J. (2016). Testing nursing students' clinical judgment in a patient deterioration simulation scenario: Development of a situation awareness instrument. *Nurse Education Today*, *38*, 61–67. https://doi.org/10.1016/j.nedt.2015.12.015

Layton, C., Smith, P. J., & McCoy, C. E. (1994). Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation. *Human Factors*, *36*(1), 94–119. https://doi.org/10.1177/001872089403600106

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Li, L., Li, Y., Zhang, Y., Xu, G., Zeng, J., & Feng, X. (2022). Formation Control of Multiple Autonomous Underwater Vehicles under Communication Delay, Packet Discreteness and Dropout. *Journal of Marine Science and Engineering*, *10*(7), Article 7. https://doi.org/10.3390/jmse10070920

Lichacz, F. M. J. (2009). Calibrating Situation Awareness and Confidence Within a Multinational Coalition Operation. *Military Psychology*, *21*(3), 412–426. https://doi.org/10.1080/08995600902914800

Loft, S., Bowden, V., Braithwaite, J., Morrell, D. B., Huf, S., & Durso, F. T. (2015). Situation awareness measures for simulated submarine track management. *Human Factors*, *57*(2), 298–310. https://doi.org/10.1177/0018720814545515

Lunghi, G., Marin, R., Di Castro, M., Masi, A., & Sanz, P. J. (2019). Multimodal Human-Robot Interface for Accessible Remote Robotic Interventions in Hazardous Environments. *IEEE Access*, *7*, 127290–127319. https://doi.org/10.1109/ACCESS.2019.2939493

Manzey, D., Luz, M., Mueller, S., Dietz, A., Meixensberger, J., & Strauss, G. (2011). Automation in Surgery: The Impact of Navigated-Control Assistance on Performance, Workload, Situation Awareness, and Acquisition of Surgical Skills. *Human Factors*, *53*(6), 584–599. https://doi.org/10.1177/0018720811426141

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human Performance Consequences of Automated Decision Aids: The Impact of Degree of Automation and System Experience. *Journal of Cognitive Engineering and Decision Making*, *6*(1), 57–87. https://doi.org/10.1177/1555343411433844

McCamy, M. B., Otero-Millan, J., Leigh, R. J., King, S. A., Schneider, R. M., Macknik, S. L., & Martinez-Conde, S. (2015). Simultaneous Recordings of Human Microsaccades and Drifts with a Contemporary Video Eye Tracker and the Search Coil Technique. *PLOS ONE*, *10*(6), e0128428. https://doi.org/10.1371/journal.pone.0128428

Merritt, S., Sinha, R., Curran, P., & Ilgen, D. (2015). Attitudinal predictors of relative reliance on human vs. Automated advisors. *International Journal of Human Factors and Ergonomics*, *3*, 327. https://doi.org/10.1504/IJHFE.2015.072982

Metzger, U., & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors*, *47*(1), 35–49. https://doi.org/10.1518/0018720053653802

Mogford, R. H. (1997). Mental Models and Situation Awareness in Air Traffic Control. *The International Journal of Aviation Psychology*, *7*(4), 331–341. https://doi.org/10.1207/s15327108ijap0704_5

Moore, K., & Gugerty, L. (2010). *Development of a Novel Measure of Situation Awareness: The Case for Eye Movement Analysis*. *54*(19), 1650–1654. https://doi.org/10.1177/154193121005401961

Morrow, D., North, R., & Wickens, C. D. (2005). Reducing and Mitigating Human Error in Medicine. *Reviews of Human Factors and Ergonomics*, *1*(1), 254–296. https://doi.org/10.1518/155723405783703019

NASA. (n.d.-a). *NASA TASK LOAD INDEX*. Retrieved June 12, 2023, from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://ntrs.nasa.gov/api/citations/20000021488/downloads/20000021488.pdf

NASA. (2022). *How Perseverance Drives on Mars*. NASA Mars Exploration. https://mars.nasa.gov/resources/26660/how-perseverance-drives-on-mars

NASA. (n.d.b). *Moving around Mars—NASA*. https://mars.nasa.gov/mer/mission/timeline/surfaceops/navigation/

Neyedli, H. F., Hollands, J. G., & Jamieson, G. A. (2011). Beyond Identity: Incorporating System Reliability Information Into an Automated Combat Identification System. *Human Factors*, *53*(4), 338–355. https://doi.org/10.1177/0018720811413767

Ni, J., Hu, J., & Xiang, C. (2021). A review for design and dynamics control of unmanned ground vehicle. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *235*(4), 1084–1100.

Nilsson, R., Gärling, T., & Lützhöft, M. (2009). An experimental simulation study of advanced decision support system for ship navigation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(3), 188–197. https://doi.org/10.1016/j.trf.2008.12.005

O'Brien, K. S., & O'Hare, D. (2007). Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*, *50*(7), 1064–1091. https://doi.org/10.1080/00140130701276640

Parasuraman, R., Cosenzo, K., & de Visser, E. (2009). Adaptive Automation for Human Supervision of Multiple Uninhabited Vehicles: Effects on Change Detection, Situation Awareness, and Mental Workload. *Military Psychology - MIL PSYCHOL*, *21*, 270–297. https://doi.org/10.1080/08995600902768800

Parasuraman, R., & Manzey, D. H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Parasuraman, R., Mouloua, M., Hillburn, B. (1999). *Adaptive Aiding and Adaptive Task Allocation Enhance Human-Machine Interaction* (M. Scerbo, M., Mouloua, Ed.).

Parasuraman, R., & Riley, V. (1997). *Humans and Automation: Use, Misuse, Disuse, Abuse*. *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans: A Publication of the IEEE Systems, Man, and Cybernetics Society*, *30*(3), 286–297. https://doi.org/10.1109/3468.844354

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160. https://doi.org/10.1518/155534308X284417

Parse, R. R. (2018). Situational Awareness: A Leadership Phenomenon. *Nursing Science Quarterly*, *31*(4), 317–318. https://doi.org/10.1177/0894318418792888

Pierce, R. (2012). The Effect of SPAM Administration During a Dynamic Simulation. *Human Factors The Journal of the Human Factors and Ergonomics Society*, *54*. https://doi.org/10.1177/0018720812439206

Psychtoolbox. (n.d.). *Psychtoolbox*. Retrieved June 11, 2023, from https://www.psychtoolbox.net/

Reichenbach, J., Onnasch, L., & Manzey, D. (2011). Human Performance Consequences of Automated Decision Aids in States of Sleep Loss. *Human Factors*, *53*(6), 717–728. https://doi.org/10.1177/0018720811418222

Rovira, E., McGarry, K., & Parasuraman, R. (2007). Effects of Imperfect Automation on Decision Making in a Simulated Command and Control Task. *Human Factors*, *49*(1), 76–87. https://doi.org/10.1518/001872007779598082

Rovira, E., & Parasuraman, R. (2010). Transitioning to Future Air Traffic Management: Effects of Imperfect Automation on Controller Attention and Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *52*(3), 411–425. https://doi.org/10.1177/0018720810375692

Sarter, N., & Woods, D. (1991). Situation Awareness: A Critical But Ill-Defined Phenomenon. *Int. J. Aviat. Psychol*, *1*. https://doi.org/10.1207/s15327108ijap0101_4

Schneider, F. E., & Wildermuth, D. (2011). An autonomous unmanned vehicle for CBRNE reconnaissance. *2011 12th International Carpathian Control Conference (ICCC)*, 347–352. https://doi.org/10.1109/CarpathianCC.2011.5945877

Shakhatreh, H., Sawalmeh, A. H., Al-Fuqaha, A., Dou, Z., Almaita, E., Khalil, I., Othman, N. S., Khreishah, A., & Guizani, M. (2019). Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges. *IEEE Access*, *7*, 48572–48634. https://doi.org/10.1109/ACCESS.2019.2909530

Sheridan, T., Verplank, W., & Brooks, T. (1978). *Human and Computer Control of Undersea Teleoperators*.

Smith, A. G., & Jamieson, G. A. (2012). Level of Automation Effects on Situation Awareness and Functional Specificity in Automation Reliance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 2113–2117. https://doi.org/10.1177/1071181312561448

SR Research. (n.d.-a). *Software Integration*. Fast, Accurate, Reliable Eye Tracking. Retrieved June 11, 2023, from https://www.sr-research.com/software-integration/

SR Research. (n.d.-b). *What are the recommended Display PC specifications for Experiment Builder?* Retrieved June 11, 2023, from https://www.sr-research.com/support/thread-84.html

SR Research. (2005). *EyeLink® II User Manual Version 2.14*.

SR Research. (2021, June 4). [Video]. https://www.youtube.com/watch?v=KJrnG4Tw-74

Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 590–605. https://doi.org/10.1016/j.trf.2018.11.006

Sternberg, S. (1966). High-speed scanning in human memory. *Science (New York, N.Y.)*, *153*(3736), 652–654. https://doi.org/10.1126/science.153.3736.652

Stoyanov, T., Vaskevicius, N., Mueller, C. A., Fromm, T., Krug, R., Tincani, V., Mojtahedzadeh, R., Kunaschk, S., Mortensen Ernits, R., Canelhas, D. R., Bonilla, M., Schwertfeger, S., Bonini, M., Halfar, H., Pathak, K., Rohde, M., Fantoni, G., Bicchi, A., Birk, A., … Echelmeyer, W. (2016). No More Heavy Lifting: Robotic Solutions to the Container Unloading Problem. *IEEE Robotics & Automation Magazine*, *23*(4), 94–106. https://doi.org/10.1109/MRA.2016.2535098

Strater, L., Faulkner, L., Hyatt, J., & Endsley, M. (2006). *Supporting Situation Awareness Under Data Overload in Command and Control Visualizations*.

Strayer, D., Cooper, J., & Drews, F. (2004). What do Drivers Fail to See When Conversing on a Cell Phone? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*. https://doi.org/10.1177/154193120404801902

Tatasciore, M., Bowden, V. K., Visser, T. A. W., Michailovs, S. I. C., & Loft, S. (2020). The Benefits and Costs of Low and High Degree of Automation. *Human Factors*, *62*(6), 874–896. https://doi.org/10.1177/0018720819867181

Tollefsen, C., Hunter, A., Randall, T., & Isenor, A. (2017). *Recommendations for implementation of the Command Reconnaissance Area Coordination and Control Environmental Network (CRACCEN) project*. https://publications.gc.ca/site/eng/9.854507/publication.html

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Vaidya, A. R., & Badre, D. (2020). Neural Systems for Memory-based Value Judgment and Decision-making. *Journal of Cognitive Neuroscience*, *32*(10), 1896–1923. https://doi.org/10.1162/jocn_a_01595

Velusamy, P., Rajendran, S., Mahendran, R. K., Naseer, S., Shafiq, M., & Choi, J.-G. (2022). Unmanned Aerial Vehicles (UAV) in Precision Agriculture: Applications and Challenges. *Energies*, *15*(1), Article 1. https://doi.org/10.3390/en15010217

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In *Automation and human performance: Theory and applications* (pp. 183–200). Lawrence Erlbaum Associates, Inc.

Wickens, C. D. (1992). *Engineering psychology and human performance, 2nd ed* (pp. xv, 560). HarperCollins Publishers.

Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, *8*, 201–212. https://doi.org/10.1080/14639220500370105

Wickens, C. D., Helton, W. S., Hollands, J. G., & Banbury, S. (2021). *Engineering psychology and human performance* (5th edition.). Routledge.

Wiener, J. M., Ehbauer, N. N., & Mallot, H. A. (2009). Planning paths to multiple targets: Memory involvement and planning heuristics in spatial problem solving. *Psychological Research PRPF*, *73*(5), 644–658. https://doi.org/10.1007/s00426-008-0181-3

Wiggins, M. W. (2021). A behaviour-based approach to the assessment of cue utilisation: Implications for situation assessment and performance. *Theoretical Issues in Ergonomics Science*, *22*(1), 46–62. https://doi.org/10.1080/1463922X.2020.1758828

Williams, K. J., Yuh, M. S., & Jain, N. (2023). A Computational Model of Coupled Human Trust and Self-Confidence Dynamics. ACM Transactions on Human-Robot Interaction. https://doi.org/10.1145/3594715

Woxenius, J. (2021). Maritime Route Planning. In R. Vickerman (Ed.), *International Encyclopedia of Transportation* (pp. 570–576). Elsevier. https://doi.org/10.1016/B978-0-08-102671-7.10485-3

Yu, Y., & Wei, J. (2009). Development on automatic conveying system for dangerous goods. *2009 International Conference on Mechatronics and Automation*, 1661–1665. https://doi.org/10.1109/ICMA.2009.5246318

Yuan, S., Li, Y., Bao, F., Xu, H., Yang, Y., Yan, Q., Zhong, S., Yin, H., Xu, J., Huang, Z., & Lin, J. (2023). Marine environmental monitoring with unmanned vehicle platforms: Present applications and future prospects. *Science of The Total Environment*, *858*, 159741. https://doi.org/10.1016/j.scitotenv.2022.159741

Zhou, Y., Wu, D., Wang, C., Sun, K., Xu, P., Wang, Z., & Xiao, W. (2022). The effect of working memory training on situation awareness in a flight simulator. *Cognition, Technology & Work*, *24*, 1–8. https://doi.org/10.1007/s10111-022-00707-1

Appendix A – Psychology Participant Pool (SONA) Contact Sheet

Study Title: The Impact of Autonomous Planning Aids in a Navigational Task
Humanities Research Ethics Board approval code: (*2021-5796)*
Principal Investigator: Dr. Heather Neyedli, School of Health and Human Performance
Co-investigator: Grace Barnhart, Masters Student

Description: The study aims to evaluate how automated aids my influence an individual's attention during planning and monitoring in a busy environment. We are interested in determining whether using an autonomous system in the planning phase of navigation task affects your ability to monitor the system. For your participation in this study, you will interact with a computer simulation. Although this is not designed to simulate any specific environment, it mimics the demands of jobs such as naval navigators, air traffic controllers, and search and rescue operators. You will be asked to wear an eye tracking device that is mounted to your head. Eye movements are collected to examine how you visually interact with the simulation and the visual attention allotted to competing demands. All together the experiment is expected to last 1.5-2 hours. The study will be conducted in the Cognitive Motor and Performance (CaMP) Lab, which is in the Dalplex at Dalhousie University. Compensation from this experiment includes $10 if you perform the task well (rank within the top 25% of participants), as well as 2 (two) SONA credits if you are currently enrolled in a psychology class at Dalhousie University. The information gained form this experiment will contribute to our understanding of human-automation interactions and factors that affect performance in navigational tasks.

Inclusion criteria: Must be between the ages of 18 and 65 years or old, normal or corrected-to-normal vision without the use of glasses.

Location: Cognitive and Motor Performance Lab, Dalplex 218C

Appendix B – Script for Recruiting Participants Using Word of Mouth

Hello, my name is __(insert name of researcher)__. I am currently conducting a study under the supervision of Dr. Heather Neyedli. The study aims to evaluate how automated aids my influence an individual's attention during planning and monitoring in a busy environment. We are interested in determining whether using an autonomous system in the planning phase of navigation task affects your ability to monitor the system. For your participation in this study you will interact with a computer simulation. Although this is not designed to simulate any specific environment, it mimics the demands of jobs such as naval navigators, air traffic controllers, search and rescue operators. You will be asked to wear an eye tracking device that is mounted to your head. Eye movements are collected to examine how you visually interact with the simulation and the attention an individual allots to competing demands. All together the experiment is expected to last 2 hours. The study will be conducted by __(insert name of researcher)__ in the Dalplex in the Cognitive Motor and Performance Lab at Dalhousie University. Compensation from this experiment includes $10 if you perform the task well, as well as 2 (two) SONA credits if you are currently enrolled in a psychology class at Dalhousie University. The information gained form this experiment will contribute to our understanding of human-automation interactions and factors that affect performance when humans are paired with autonomous aids. To be eligible to participate you must be between the ages of 18 and 65 years or old, normal or corrected-to-normal vision without the use of glasses. If you are interested in participating in this study, please contact me at __(insert email of researcher)__.

DALHOUSIE UNIVERSITY

Appendix C – Informed Consent Form

Study Title: The Impact of Autonomous Planning Aids in a Navigational Task

Lead Researcher:
Dr. Heather Neyedli
Kinesiology, Faculty of Health
Department of Psychology
Dalhousie University
Phone: (902) 494-6786
Email: hneyedli@dal.ca

Other Researchers:
Grace Barnhart
Masters Student
Division of Kinesiology, Health and Human Performance, Dalhousie University
Email: grace.barnhart@dal.ca

Affiliated Researchers:
Dr. David Westwood – Kinesiology – Dalhousie University
Dr. Aren Hunter – Atlantic Research Center – Defense Research and Development Canada

Funding:
This project has received funding from Department of National Defense - Innovation for Defence Excellence and Security (IDEaS).

Introduction
You are invited to take part in the research project described below. The Social Sciences and Humanities Research Ethics Board of Dalhousie University has reviewed the project and found it to conform to current ethical guidelines. These guidelines require:

1) That you be informed of the purpose of the research project and any possible inconveniences, risks, or benefits.
2) That the character of the task required be explained to you
3) That you understand that participation is voluntary, and that you may decline to continue participation at any point throughout the course of the research project, without loss of expected compensation, nor any academic impact as a result of deciding whether or not to participate.
4) That you be assured that all information assembled is entirely confidential.

Should you have any further question after reading this informed consent form, please feel free to ask question about anything that may have been unclear.

Purpose of the Study
The purpose of this study is to 1) determine how the level of automation in a route planning task affects situational awareness, workload, and one's ability to later monitor the task and 2) to determine whether the use of autonomous aids is related to changes in visual eye patterns.

Who Can Take Part in the Research Study
You are eligible to participate in this study if you are between the ages of 18 and 65 years old and have normal or correct-to-normal vision without glasses.

What You Will Be Asked to Do
The experiment consists of a computer simulation of unmanned vehicles as they explore a region where some areas are known to be safe, some are known to be unsafe, and the status of others is unknown. In the simulation, you will take on the role as their operator and will monitor their movement to ensure they don't enter into unsafe or unknown areas. You will simultaneously complete a search task where you will try to locate and click on a "T" among an array of "L's".

The experiment is expected to take a total of 1.5-2 hours. First, the experimenter will provide you with detailed instructions about the task. After the instructions you will have the opportunity to ask the researcher about any questions or concerns you may have regarding the experiment or the simulation.

Following explanation of the experiment, you will complete a series of training and experimental blocks lasting between 2 – 25 minutes each. In training, the first training block you will be asked to locate and click on a target. In the second training block you will monitor the unmanned vehicles as they travel and cancel their routes if they are about to enter an unsafe area. Once familiarized with these two tasks, you will complete them simultaneously in the third training block. The experimental blocks will also combine these skills to complete them simultaneously. After each session you will be provided an opportunity to take a break.

One training is complete, <u>you will be asked to wear a head mounted eye tracker</u> for the remainder of the study. The eye tracker is worn with bands around and overtop the head, and two arms containing cameras that extend just below eye level. The system uses image processing algorithms to determine where on the screen you are looking throughout the study. During the experimental blocks, you will also be asked questions about the current environment to measure your situational awareness. Immediately after each of the experimental blocks you will be asked to complete a questionnaire measuring your workload during the block.

Possible Benefits
Participation in this study may not benefit you directly, but this study will contribute to knowledge within the field of cognitive ergonomics as well as human-computer interactions.

Compensation / Reimbursement
If you achieve a high level of performance in the simulation, you will receive $10 for participating in this study. Performance is based on a combination of correct responses in a navigational monitoring task and a visual search task. The top 25% of participants will receive a $10 bonus. Participants who are students, are registered in a psychology class at Dalhousie University and who signed up through SONA will be granted 2 (two) credit points regardless of their performance. If received, the $10 bonus is awarded in addition to credit points.

Possible Risks and Discomforts
Possible risks of participation in this study includes fatigue that may be caused by the mental effort required to perform the task in the experiment. This could lead to stress similar to playing a more challenging level on a video game. You may find the eye tracker uncomfortable. There will be frequent breaks but if you are experiencing discomfort, please let the researcher know and we will adjust the headbands.

If You Decide to Stop Participating
You may choose not to continue your participation in the study at any time. If you decide not to take part in the study or if you leave the session early, your data will be automatically withdrawn from the study. Further, you may choose to withdraw your data after you have participated. However, once data has been analyzed, it will no longer be possible to withdraw from the study. We will hold off analyzing data for 1 week following collection from the final participant (i.e. study completion) to allow you to withdraw your data after you have participated.

How Your Information Will Be Protected
Every effort to protect your privacy will be made. No identifying information will be included in publications or presentations. Minimal information about you will be collected by the research team, ensuring only required information (such as age, and information from study questionnaires) is collected.

Confidentiality: In order to protect your privacy and keep your participation in the study confidential, you will be de-identified using a study code. For the purpose of data analyses, all participants will only be identified by their study code (e.g., P001) and all data will be stored on password protected computers and spreadsheets. You will be permitted to withdraw your data up to a week after completion of the study (i.e., one week after the final participant is collected), as data typically is analyzed one week after completion. However, once data is analyzed (i.e., after one week of study completion), it will not be possible to withdraw your data.

Data Retention: information that you provide to us will be kept private. Only the researchers will have access to this information. Only anonymized data will be sent to the research team through password protected files. Your name will not appear on any of these files. We will describe and share our findings in theses, presentations, public media, journal articles, etc. This means that you will not be identified in any way in our reports. The people who work with us have an obligation to keep all research information private.

Also, we will use a participant number (not your name) in our written and computer records so that the information we have about you contains no names. All your identifying information contained on the consent will be securely stored separately from your data in a locked cabinet in the Cognitive and Motor Performance Lab.

Questions

We are happy to talk with you about any questions or concerns you may have about your participation in this research study. For further information you may contact the principal or co-investigator (information provided on the front page). If you have any ethical concerns about your participation in this research, you may also contact Research Ethics, Dalhousie University at (902) 494-1462, or email: ethics@dal.ca and reference REB file 2021-5796.

DALHOUSIE
UNIVERSITY

INFORMED CONSENT SIGNATURE PAGE

Study Title: The Impact of Autonomous Planning Aids in a Navigational Task

I have read the informed consent form and meet the requirements for participation as outlined on the screening form for this study. I have been given the opportunity to discuss the study and my questions have been answered to my satisfaction.

I agree that my study information may be used as described in this consent form.

I understand that my participation in this study is voluntary and that I may withdraw my consent from the study at any time, without penalty.

______________________________ ____________________________ ___________

Name (Please Print) Signature Date

The results of this study can be distributed by email when available. If you wish to receive the results, please indicate your preferred email address.

Email: ____

Appendix D - SONA Consent Signature Page

DALHOUSIE UNIVERSITY

SIGNATURE PAGE

(this page must be printed on a separate sheet)

- **Participants Must Read And Sign This Form To Confirm That They Understand And Accept Conditions Before Experiment Can Begin**
- **Participants Must Be Given A Copy Of This Form For Their Information And Records**

Feel free to address any questions you may have about the study to the Principal Investigator / Researcher either now, or after you have participated.

Study Title The Impact of Autonomous Planning Aids in a Navigational Task

Name of Principal Investigator Dr. Heather Neyedli

Research Supervisor (if different from PI)

Contact Person (if different from PI) Grace Barnhart

Address 6260 South Street

Telephone 902-718-6977

Email grace.barnhart@dal.ca

Psychology Department Subject Pool Policy

Individuals with specific ethical concerns should contact either the Research Supervisor or a member of the Human Research Participants & Ethics Committee of the Department of Psychology & Neuroscience, Tel: 902.494.1580, email psych.ethics@dal.ca.

Please sign below to confirm that you have had your questions answered to your satisfaction, that you are aware that all records are entirely confidential and that you may discontinue participation at any point in the study.

If you anticipate receiving educational credit points for assisting in this research, you may choose to do so as either a **Research Participant** or as an **Observer**.

If you choose to be a Research Participant, the researcher will keep your data and use it in the research project.

If you choose to be an Observer, the researcher will destroy any data that you may have provided, after you complete the study.

Please check one box below to indicate whether you choose to be a Research Participant or an Observer.

☐ **Research Participant (Use my data)**

☐ **Observer (Destroy my data)**

Participant's Signature: Date:

Researcher's Signature: Date:

FACULTY OF SCIENCE | *Department of Psychology & Neuroscience* | Life Sciences Centre |1355 Oxford Street | PO Box 15000 | Halifax NS B3H 4R2 Canada | Tel. 902.494.1580 | Fax. 902.494.6585 | psych.ethics@dal.ca | dal.ca/psychandneuro
DAL.CA

Appendix E – Demographic Information

The following information is being collected to get a general idea of the sample group that took part in this study. Please fill out this information if you are comfortable disclosing it.

Gender: _______________________ Age: ____________

Appendix F - Training and Experimental Scripts

Each script will be read by participants (or, if necessary, read to the participant) prior to each respective block.

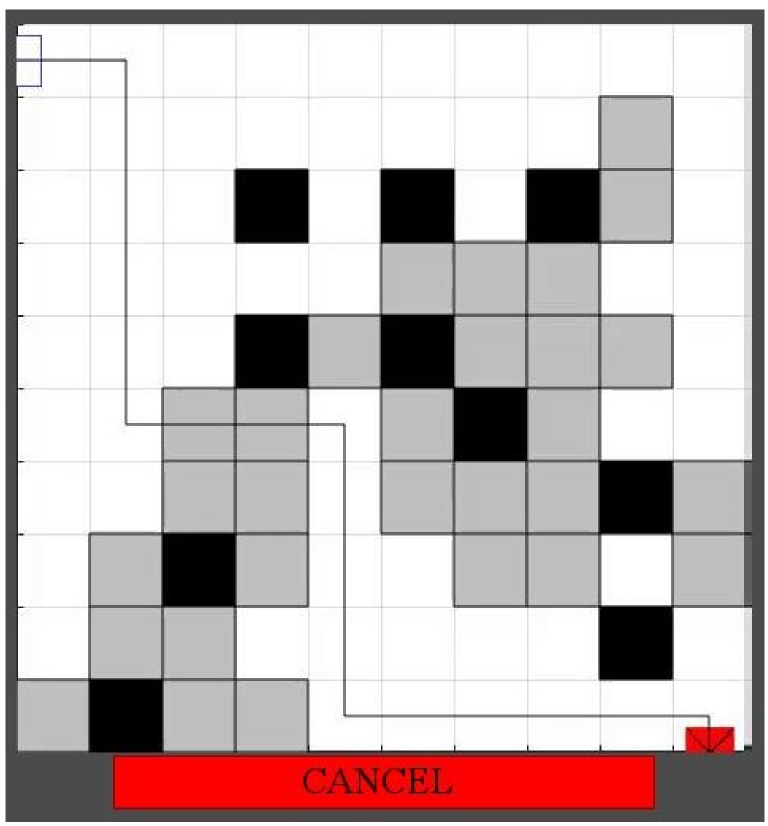
Training Block One
In this simulation, you will act as an operator of uncrewed semi-autonomous vehicles exploring remote regions in small segments. You will complete three short training blocks (2-5 minutes each) that will introduce you to the task, and then three longer experimental blocks (15-20 minutes each). In this training block you will be familiarized with a search task that has been designed to mimic duties that an operator may have to complete while monitoring an autonomous system in an applied work setting (e.g., identifying underwater mines in naval navigation, or new aircraft entering the air space in air traffic control).Your goal is to locate and click on a target shaped as a "T" among distractors in the shape of an "L" before the five second counter, displayed at the top, runs out. Here is an example of what the search task will look like.

Once the timer has elapsed, a new array and target will be presented and the timer will restart. Correct clicks within the timer duration are considered a success regardless of how quickly you click on the target. There is no feedback onscreen [feedback was provided by the researcher].

The target “T” may be presented in any orientation. All clicks within a small radius of the target are deemed to be correct, so you don’t have to click directly on the "T" but you should try to be as accurate as possible.

This block consists of ten consecutive trials and should take approximately one minute to complete. Performance here will not count towards your overall study performance score. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Training Block Two
In this training block you will be introduced to the uncrewed semi-autonomous vehicles and your role as their operator. The vehicles will be travelling through segments of a region with areas that are known to be safe or unsafe to pass through, and areas where the status is unknown. The goal of this task is to identify safe routes through selected segments of the region. The segments are represented on a grid (see below image) where the status of each grid square is indicated by their color: a black square indicates known unsafe areas that cannot be traversed, white indicates known safe areas, and grey indicates an area where the status is currently unknown. To determine safe routes through segments with areas of unknown status, you will be responsible for the vehicles as they travel from their current location (white square with a blue outline in the below image), to an end point (red square with black “V”). Here is an example of a grid segment.



The vehicles are equipped with sensors that can assess areas as they are approached and send the data back to this computer. If enough data is gathered to indicate that an unknown area is safe, it will be reflected on the map by the grey squares changing to white; otherwise, if the vehicles don't gather enough data, the map will remain unchanged. Since they are uncrewed, the vehicles require instructions (e.g., route to travel, when to start) which are sent remotely from this computer.

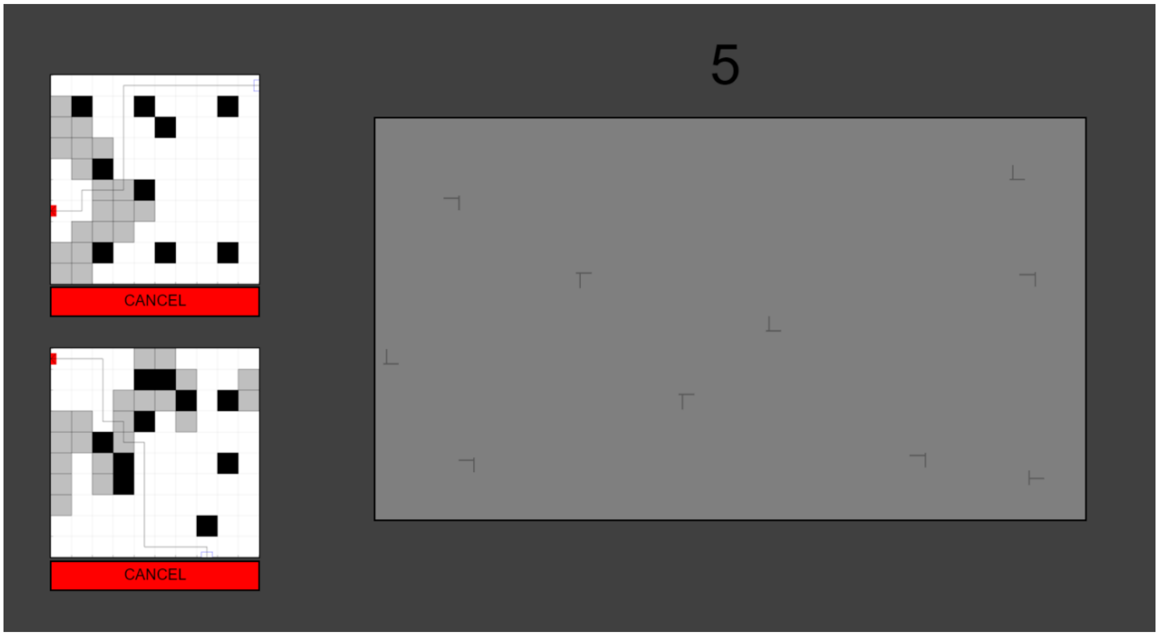
To facilitate this task, you will be assisted by an autonomous planning aid that will compute a route between the start and end point for each segment and send the necessary instructions to the vehicles. The planning aid is programmed to minimize the distance travelled while avoiding unsafe areas and minimizing travel through areas of unknown status. For each segment, the aid will display its chosen route for you to review. You may take up to 30 second to review the route. If you would like to proceed before the time has elapsed, you may click the red “start” button. Once you have reviewed the route, a message will be sent from the computer to the vehicle directing it to begin moving and to follow the outlined route. As the vehicle travels along the route, its location (represented by a blue square) will be updated on your monitor.

As the operator your job is to monitor the vehicles and ensure they don’t cross into unsafe or unknown areas so they are not lost or damaged. If they are about to enter an unknown area, their route should be cancelled by clicking the red “CANCEL” button. Clicking this will send a message to the vehicles to stop. The vehicles used here have a range of about two grid squares but need some time to collect and analyse sensor data before an area can be determined safe. If the vehicle is within one grid square of an unknown area their route should be cancelled. You should not cancel the route too early (i.e., before the one square range), because it will not have had the opportunity to evaluate the status of the area.

To familiarize you with the task we will show you recorded data from ten previously explored segments that have a high ratio of segments that require intervention. In the first five segments, unknown areas will be updated to be safe as they are approached by the vehicle. You should watch these and let the vehicle complete the route. The last five contain areas not found to be safe, so the vehicle should be stopped before it enters into the unknown areas. In this training block a failed trial occurs if you cancel the route before it is within one grid square of an unknown status area, or if the vehicle travels into a grey unknown status area. A successful trial occurs if the route is navigated without entering an area of unknown status, or if the route is cancelled within one grid square of an unknown status area. Please note that for training purposes we are intentionally showing you a high ratio of segments with unknown areas that are not deemed to be safe. In the segments you will be examining after your training, most of the areas of unknown status are expected to be safe.

Please take as much time as you would like to review these instructions and direct questions to the researcher. The block consists of ten trials and should take you five to ten minutes to complete.

Training Block Three
This training block will combine the two tasks you have already practiced. You will maintain the role of an operator of uninhabited vehicles that are exploring routes through regions with unknown status to determine safe routes. Since many of the segments in this region are expected to be safe, you will now monitor two vehicles as they explore unique segments while simultaneously completing the previously practiced search task. Please see the below image for an example of how the tasks will be displayed.

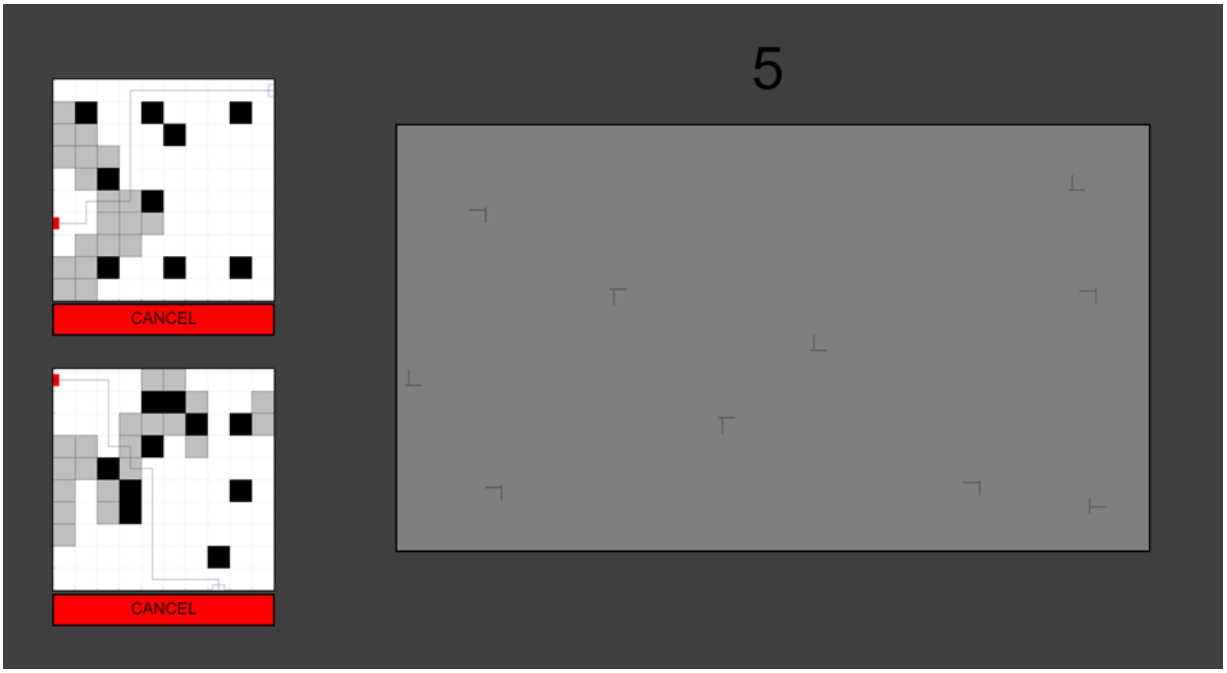


Your primary task will be to locate and click on the target "T" in the search task. As practiced, correct clicks within the five second timer are considered a success regardless of how quickly you click on the target. Once the five second time timer has elapsed, a new array and target will be presented, and the timer will restart.

For the navigation tasks on the left, you will still be assisted by an autonomous planning aid. You will have the opportunity to review the segments for a maximum of one minute, or can click "START" to proceed. The review time has been doubled from the 30 seconds allotted in the previous block as you will be responsible for reviewing two routes in separate segments for each trial. Once the review is complete, the computer will send a message to both vehicles, and they will begin travelling their respective routes. At the same time, the search task will commence on the right side of the screen. While your primary goal the search task, it is also important to ensure neither vehicle travels into an unknown area where they may be lost or damaged.

This training block consists of five trials and should take you five to ten minutes to complete. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Experimental Block One
You will once again take on the role of an operator for uninhabited vehicles that are exploring routes through areas with unknown status to evaluate routes. As practiced in the previous block, you will monitor two vehicles as they explore unique segments while simultaneously completing a search task. Please see the below image for a reminder of how the tasks will be displayed.



Your primary task will still be to locate and click on the target "T" in the search task. As practiced, correct clicks within the five second timer are considered a success regardless of how quickly you click on the target. Once the five second time timer has elapsed, a new array and target will be presented, and the timer will restart. For the navigation tasks on the left, you will still be assisted by an autonomous planning aid. You will have the opportunity to review the segments for a maximum of one minute , or can click "START" to proceed. Once the review is complete, the computer will send a message to both vehicles, and they will begin travelling their respective routes. At the same time, the search task will commence on the right side of the screen. While your primary goal the search task, it is also important to ensure neither vehicle travels into an unknown area where they may be lost or damaged.

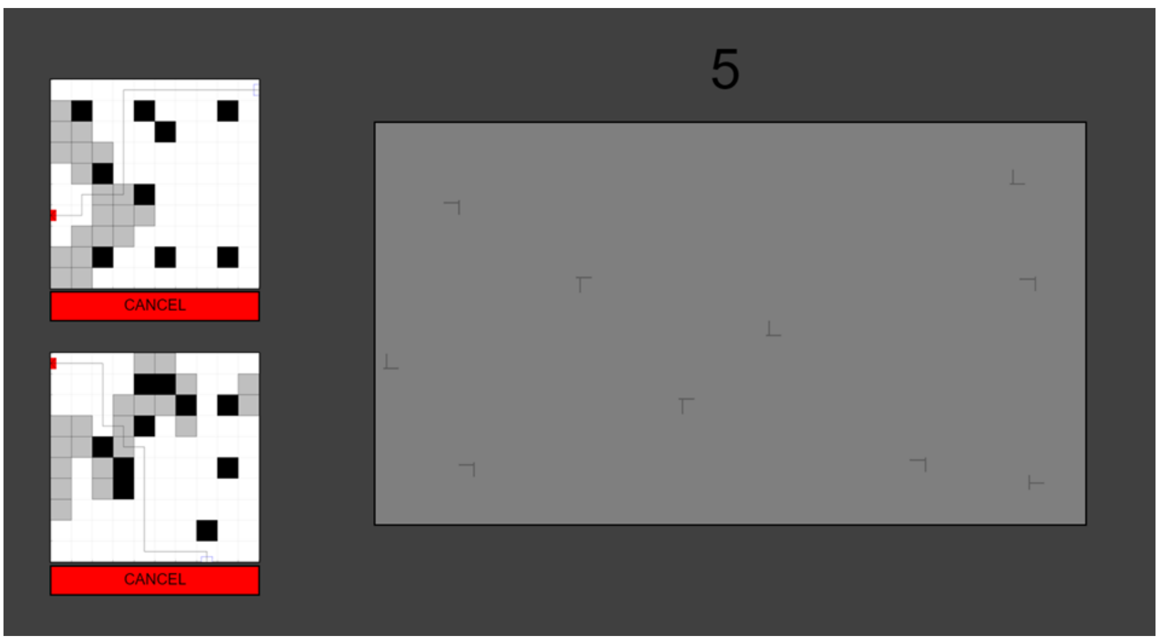
During the trials, the simulation may randomly pause and direct you to the adjacent monitor. When directed, please use the monitor and mouse to your left and follow the prompts on screen. You will be answering questions about the state of the simulation just before the screen was paused. For example, the questions might ask you about the location of a given vehicle or which direction it was travelling. Answers to these questions to not count towards your performance score.

Your performance in this, and subsequant blocks will be used to determine your overall performance for the study. You will earn one point for each target correctly located in the search task. However, if the vehicle enters into an area that is not known to be safe (i.e., a grey or black square) you will lose all points earned for that trial. All scores will be calculated after data collection and are not available throughout the study.

People that perform in the top 25% of all participants will be sent $10 through the email address provided once data collection is completed. If you registered for this study through SONA you will receive the SONA credit regardless of your performance, and any money earned willl be in addition to the SONA points.

The block consists of twenty-two trials and should take you fifteen to twenty minutes to complete. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Experimental Block Two – low LOA
In this block you will maintain your role as an operator of semi-autonomous uninhabited vehicles. As in the previous block, you will monitor two vehicles as they explore unique segments while simultaneously completing the search task. Please see the below image for a reminder of how the tasks will be displayed.



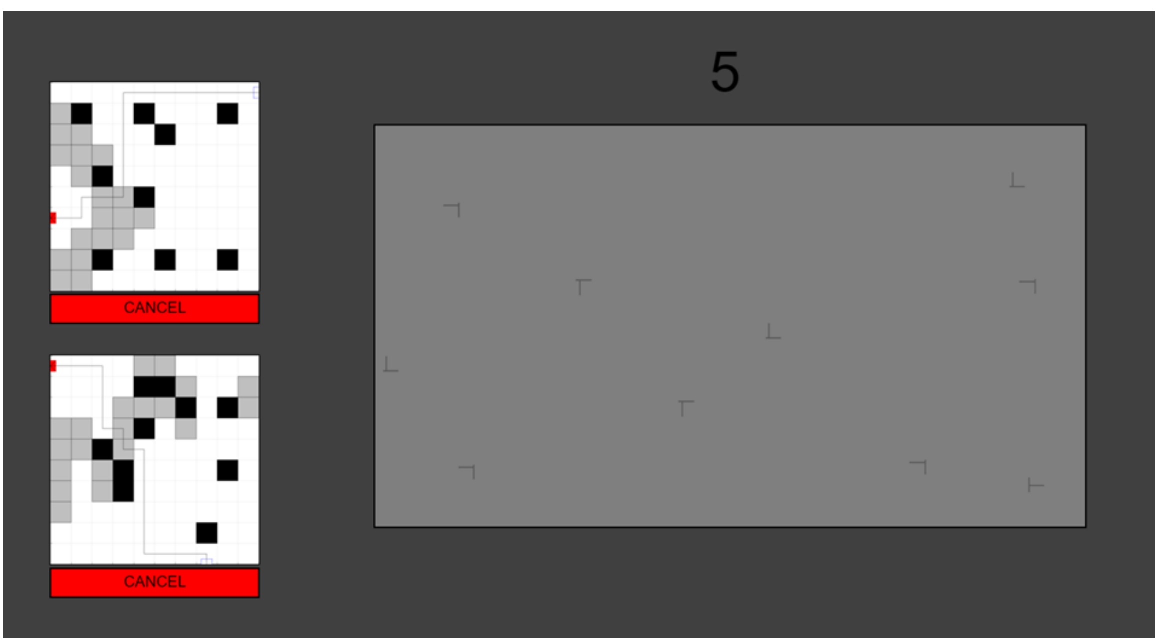
Your primary task will still be to locate and click on the target "T" in the search task. As practiced, correct clicks within the five second timer are considered a success regardless of how quickly you click on the target. Once the five second timer has elapsed, a new array and target will be presented, and the timer will restart.

You will still be assisted by a planning aid, however, in this block, you are also asked to input your evaluation of the environment by recommending a route before reviewing the autonomously planned route. This planning exercise is completed to evaluate if or how your diagnosis varies from that of the automated system. Before each trial you will be prompted to complete a planning exercise on the monitor to your left. Prompts will walk you through the exercise and will ask you to map a route from the vehicles current point (white square with blue outline) to the end point (red square with "V"). The route should limit the route distance while ensuring that it doesn't pass through any unsafe regions (black) and minimizes the time spent in areas of unknown status (grey). There is no need to explore the entire area, please create a direct route from the current location to the end point. The vehicle can only move either vertically or horizontally (it cannot move on a diagonal). You may take up to twenty seconds. Once you have completed the first route, you will be prompter to complete two more that are each unique. When three routes for the top segment are completed, the bottom segment will appear. Please repeat these steps. Once three routes are completed for each segment, a message will direct you back to this monitor where you will follow the same sequence as the previous block.

During the trial the simulation may randomly pause and direct you to the adjacent monitor. When directed, please use the monitor and mouse to your left and follow the prompts on screen. You will be answering questions about the state of the simulation just before the screen was paused. Answers to these questions do not count towards your performance score. The block consists of twenty-two trials and should take you twenty-five to thirty minutes to complete. Performance in this block will count towards your overall performance score in this study. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Experimental Block Two – moderate LOA

In this block you will maintain your role as an operator of semi-autonomous uninhabited vehicles. As in the previous block, you will monitor two vehicles as they explore unique segments while simultaneously completing the search task. Please see the below image for a reminder of how the tasks will be displayed.



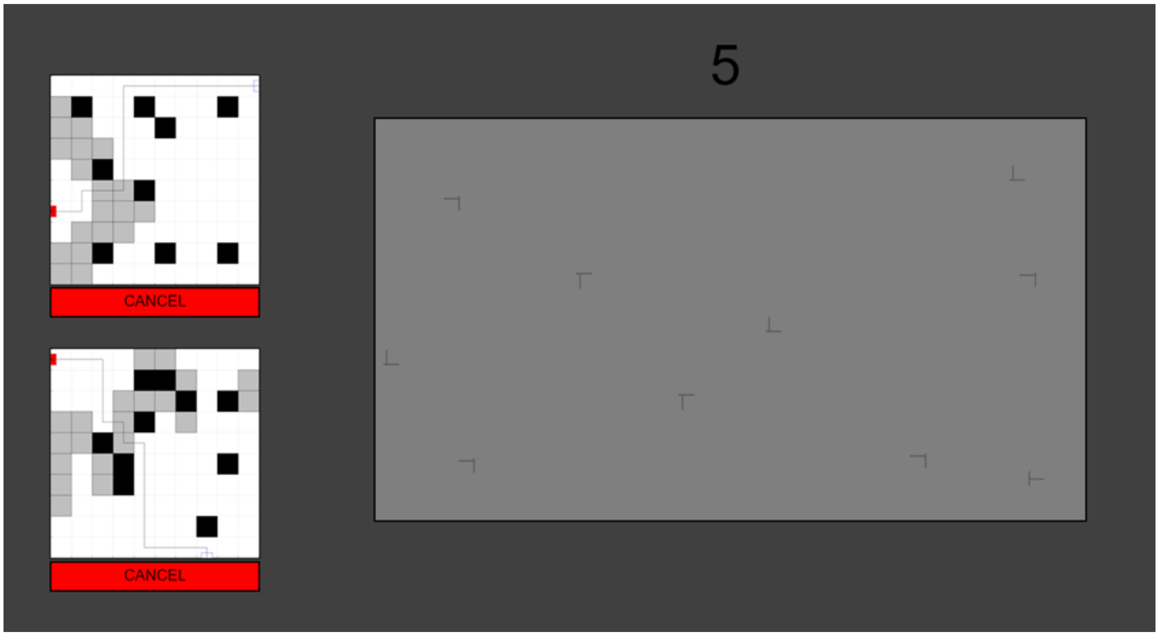
Your primary task will still be to locate and click on the target "T" in the search task. As practiced, correct clicks within the five second timer are considered a success regardless of how quickly you click on the target. Once the five second timer has elapsed, a new array and target will be presented and the timer will restart.

You will still be assisted by a planning aid, however, in this block, you are also asked to input your evaluation of the environment by ranking three potential routes, that have been generated by the planning aid, from best to worst. Before each trial you will be prompted to complete this exercise on the monitor to your left. Prompts will walk you through the exercise and will ask you to select your first, second and third choice route in turn. The route should limit the route distance while ensuring that it doesn't pass through unsafe regions (black squares) and minimized the time spent in areas with unknow status (grey squares). It should take you about twenty seconds to register responses for each segment. Once completed, a message will direct you back to this monitor where you will follow the same sequence as the previous block.

During the trial the simulation may randomly pause and direct you to the adjacent monitor. When directed, please use the monitor and mouse to your left and follow the prompts on screen. You will be answering questions about the state of the simulation just before the screen was paused. Answers to these questions do not count towards your performance score. The block consists of twenty-two trials and should take you twenty-five to thirty to complete. Performance in this block will count towards your overall performance score in this study. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Experimental Block Two – high LOA

In this block you will maintain your role as an operator of semi-autonomous uninhabited vehicles. As in the previous block, you will monitor two vehicles as they explore unique segments while simultaneously completing the search task. Please see the below image for a reminder of how the tasks will be displayed.

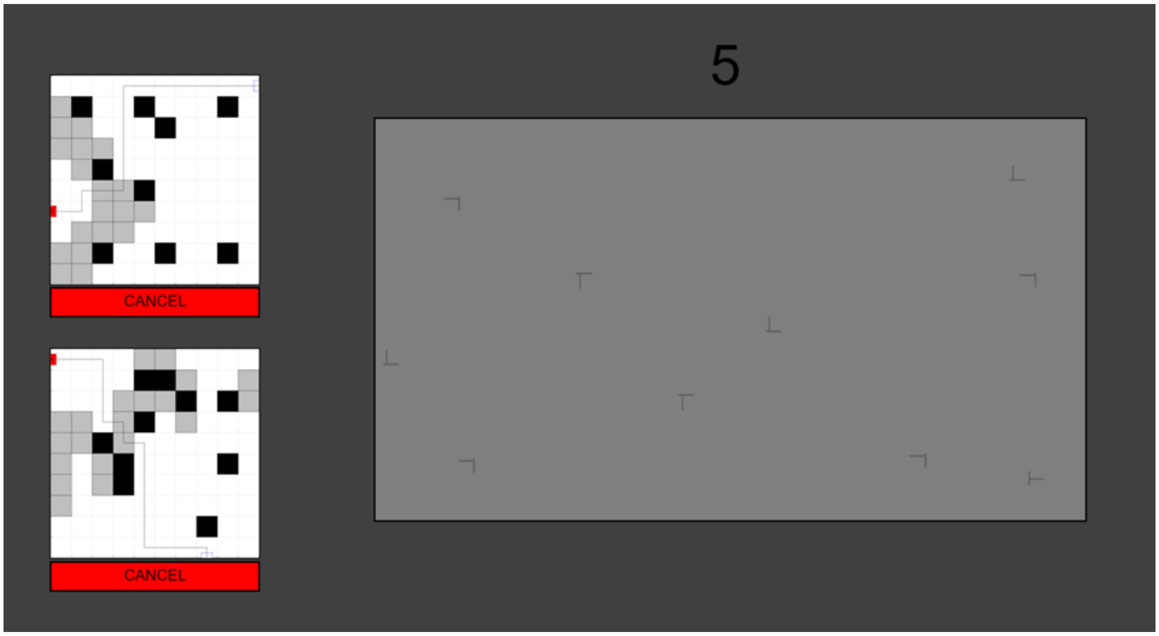


Your primary task will still be to locate and click on the target “T” in the search task. As practiced, correct clicks within the five second timer are considered a success regardless of how quickly you click on the target. Once the five second timer has elapsed, a new array and target will be presented and the timer will restart.

You will still be assisted by a planning aid, however, in this block, you will review each route separately for twenty seconds. The routes should limit the route distance while ensuring that it doesn’t pass through any unsafe regions (black squares) and minimize the time spent in the areas of unknown status (grey squares). Once completed, a message will direct you back to this monitor where you will follow the same sequence as the previous block.

During the trial the simulation may randomly pause and direct you to the adjacent monitor. When directed, please use the monitor and mouse to your left and follow the prompts on screen. You will be answering questions about the state of the simulation just before the screen was paused. Answers to these questions do not count towards your performance score. The block consists of twenty-two trials and should take you twenty-five to thirty minutes to complete. Performance in this block will count towards your overall performance score in this study. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Experimental Block Three
This block will run similar to the first block where planning exercises were not required. You will continue to act as an operator for uninhabited vehicles that are exploring routes through unknown regions to evaluate routes. You will monitor two vehicles as they explore unique segments while simultaneously completing a search task. Please see the below image for a reminder of how the tasks will be displayed.



Your primary task will still be to locate and click on the target “T” in the search task. As practiced, correct clicks within the five second timer are considered a success regardless of how quickly you click on the target. Once the five second timer has elapsed, a new array and target will be presented and the timer will restart.

For the navigation tasks on the left, you will still be assisted by an autonomous planning aid. You will have the opportunity to review the segments for a maximum of one minute, or can click “START” to proceed. Once the review is complete, the computer will send a message to both vehicles, and they will begin travelling their respective routes. At the same time, the search task will commence on the right side of the screen. While your primary goal the search task, it is also important to ensure neither vehicle travels into an unknown area where they may be lost or damaged.

During the trial the simulation may randomly pause and direct you to the adjacent monitor. When directed, please use the monitor and mouse to your left and follow the prompts on screen. You will be answering questions about the state of the simulation just before the screen was paused. Answers to these questions do not count towards your performance score. This block consists of twenty-two trials and should take you fifteen to twenty minutes to complete. Performance in this block will count towards your overall performance score in this study. Please take as much time as you would like to review these instructions and direct questions to the researcher.

Appendix G – NASA Task Load Index

Note. Participants responded to each question on a 10-point Likert scale.

1. How mentally demanding was the task?

Very low									**Very high**
1	2	3	4	5	6	7	8	9	10
O	O	O	O	O	O	O	O	O	O

2. How physically demanding was the task?

Very low									**Very high**
1	2	3	4	5	6	7	8	9	10
O	O	O	O	O	O	O	O	O	O

3. How hurried or rushed was the pace of the task?

Very low									**Very high**
1	2	3	4	5	6	7	8	9	10
O	O	O	O	O	O	O	O	O	O

4. How successful were you in accomplishing what you were asked to do?

Failure									**Perfect**
1	2	3	4	5	6	7	8	9	10
O	O	O	O	O	O	O	O	O	O

5. How hard did you have to work to accomplish you level of performance?

Very low									**Very high**
1	2	3	4	5	6	7	8	9	10
O	O	O	O	O	O	O	O	O	O

6. How insecure, discourages, irritated, stressed, and annoyed were you?

Very low									**Very high**
1	2	3	4	5	6	7	8	9	10
O	O	O	O	O	O	O	O	O	O

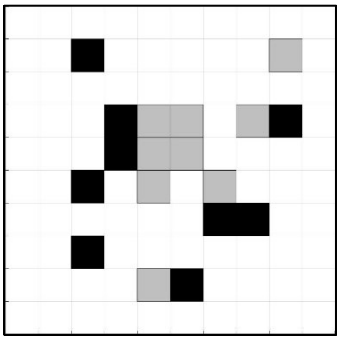
NASA TLX Rating scale definitions. To be used in conjunction with the above rating scale.

Scale Title	**Endpoints**	**Description**
MENTAL DEMAND	*Low/High*	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task east or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	*Low/High*	How much physical activity was required (e.g., pushing, pulling, turning, controlling, etc.)? Was the task demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	*Low/High*	How much time pressure did you fell due to the rate or pace at which the tasks or elements occurred? Was the pace slow and leisure or rapid and frantic?
EFFORT	*Low/High*	How hard did you have to work (mental and physical) to accomplish your level of performance?
PERFORMANCE	*Good/Poor*	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing this task?
FRUSTRATION LEVEL	*Low/High*	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

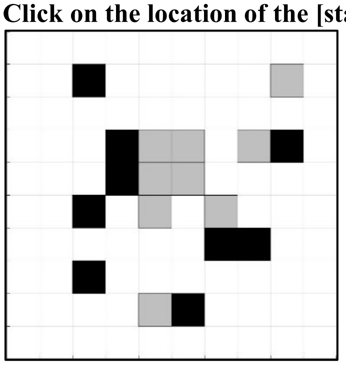
Appendix H – SAGAT Questions

Note. Text in square brackets varied by query iteration.

1. Click on the location of the vehicle in the [top/bottom] segment.



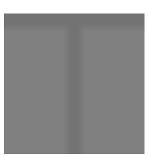
2. Click on the location of the [start/end] point in the [top/bottom] segment.

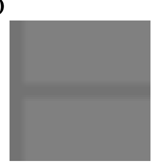


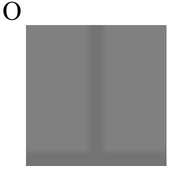
3. Which Direction is the vehicle moving in the [top/bottom] segment?

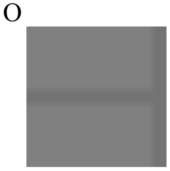
- ○ Up
- ○ Down
- ○ Left
- ○ Right
- ○ Not applicable (e.g., stopped)
- ○ I don't know

4. Select the orientation of the target in the current search task array.

O 

O 

O 

O 

O I don't know

5. In the [top/bottom] segment, in the next two grid squares of the vehicles route is there an area of unknown status?

- ○ Yes
- ○ No
- ○ I don't know
- ○ Not applicable

6. In the [top/bottom] segment, has the vehicle already passed through an area with a status that was previously not known?

- ○ Yes
- ○ No
- ○ I don't know

7. In the [top/bottom] segment, does the vehicle have to pass through a currently unknown area before the end of this trial?

- ○ Yes
- ○ No
- ○ I don't know
- ○ Not applicable

8. In the [top/bottom] segment, does the vehicle have to make a turn before the end of the trial

- ○ Yes
- ○ No
- ○ I don't know
- ○ Not applicable

9. If both routes clear, which vehicle will complete the segment first?

- ○ Top segment
- ○ Bottom segment
- ○ I don't know
- ○ Not applicable

10. If all areas of unknown status are NOT cleared, which vehicles route will have to be cancelled first?

- ○ Top segment
- ○ Bottom segment
- ○ I don't know
- ○ Not applicable

11. In the [top/bottom] segment, what is the primary direction of the vehicles route for the remainder of the trial?

- ○ Upwards
- ○ Downwards
- ○ Left
- ○ Right
- ○ Not applicable (routes cancelled)
- ○ No primary direction. Must move the same distance in more than one direction
- ○ I don't know

Appendix I – Debriefing Form

Study Title: The Impact of Autonomous Planning Aids in a Navigational Task

Researchers: Heather Neyedli and Grace Barnhart

Contact Information: grace.barnhart@dal.ca

Debriefing

The purpose of this experiment is to better understand how autonomous planning aids affect an operator's monitoring performance. The experiment consisted of three groups of participants. One group had participants complete a manual planning exercise in the second experimental block, another group had a moderately autonomous aid for the second block which displayed options for them to rank, and the third group had a highly autonomous aid for all three experimental blocks. In all groups the vehicle navigated an autonomously chosen route. Automated route planning systems are used in a variety of fields (such as air traffic control, maritime navigation, and automobile navigation) to reduce operator workload. Previous research has suggested that highly autonomous planning aids might reduce an operator's situational awareness as well as their ability to detect and respond to changes in the environment. If an optimal level of automation can be determined to reduce workload without impairing SA, performance may increase leading to a reduction in the occurrence of errors. Additionally, the act of placing value on options might be a critical process. Eye metrics were collected to evaluate whether the differences in situation awareness and performance across autonomous aids may be due to different visual strategies.

We told you at the start of the study that you would receive $10 compensation based on performance; however, all participants regardless of performance receive $10. We did this to increase your motivation on the task to more realistically mimic the motivation that an operator might have.

Questions

If after you leave today, you have questions or concerns about your participation in this research study please contact Dr. Heather Neyedli (902-494-6786, hneyedli@dal.ca) at any time with questions, comments, or concerns about the research study.

If you have any ethical concerns about your participation in this research, you may also contact Research Ethics, Dalhousie University at 902-494-1462, or email: ethics@dal.ca (*reference REB file 2021-5796)*.

Withdrawal of Data

As we stated in the consent form, you still have the opportunity to withdraw your data up until the point it is entered into analysis (typically 1 week after study completion). If you wish to do so after reading this debrief form, please inform the experimenter verbally or contact Dr. Heather Neyedli (hneyedli@dal.ca, 902-494-6786) at a later date.