2017

The Impact of Automation and Stress on Human Performance in UAV Operation

Jinchao Lin
University of Central Florida

THE IMPACT OF AUTOMATION AND STRESS ON HUMAN PERFORMANCE IN UAV OPERATION

by

JINCHAO LIN
B.S. University of Jinan, 2011
M.S. University of Central Florida, 2014

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2017

Major Professor: Gerald Matthews

ABSTRACT

The United States Air Force (USAF) has increasing needs for unmanned aerial vehicle (UAV) operators. Automation may enable a single operator to manage multiple UAVs at the same time. Multi-UAV operation may require a unique set of skills and the need for new operators calls for targeting new populations for recruitment. The objective of this research is to develop a simulation environment for studying the role of individual differences in UAV operation under different task configurations and investigate predictors of performance and stress. Primarily, the study examined the impact of levels of automation (LOAs), as well as task demands, on task performance, stress and operator reliance on automation. Two intermediate LOAs were employed for two surveillance tasks included in the simulation of UAV operation. Task demand was manipulated via the high and low frequency of events associated with additional tasks included in the simulation. The task demand and LOA manipulations influenced task performance generally as expected. The task demand manipulations elicited higher subjective distress and workload. LOAs did not affect operator workload but affected reliance behavior. Also, this study examined the role of individual differences in simulated UAV operation. A variety of individual difference factors were associated with task performance and with subjective stress response. Video gaming experience was linked to lower distress and better performance, suggesting possible transfer of skills. Some gender differences were revealed in stress response, task performance, but all the gender effects became insignificant with gaming experience controlled. Generally, the effects of personality were consistent with previous studies, except some novel findings with the performance metrics. Additionally, task demand was found to moderate the influence of personality factors on stress response and performance metrics. Specifically, conscientiousness was associated with higher subjective engagement and performance when demands were higher. This study supports future research which aims to improve the dynamic interfaces in UAV operation, optimize operator reliance on automation, and identify individuals with the highest aptitude for multi-UAV control.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) in collaboration with Air Force Research Laboratory (AFRL). The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of AFRL or the U.S. Government.

This research would not have been possible without the help, advice, and support of many. I would like to express the deepest gratitude and appreciation to my committee chair: Dr. Gerald Matthews, for serving as my primary advisor through my graduate career. Without his guidance and persistent help, this dissertation would not have been possible. I also would like to express my sincere appreciation to the other members of my committee: Dr. Lauren Reinerman-Jones, Dr. James Szalma, and Dr. Gregory Funke, for their interest, guidance, enthusiasm, and spending their valuable time evaluating my dissertation. I am grateful to AFOSR scientists for their sponsorship and support in realizing this project including Dr. Gloria Calhoun and Mr. Heath Ruff. In addition, I wish to thank Dr. C.-Y. Peter Chiu, for his help and counsel at the University of Cincinnati. I would also like to thank my colleagues at the Institute for Simulation and Training (IST) including Dr. Ryan Wohleber, for his insight, project support, and data collection efforts, Dr. Daniel Barber and Dr. Jonathan Harris, for their technical expertise and troubleshooting, and Dr. Grace Teo, for her valuable suggestions and help. Additionally, my thanks must also be extended to Ms. Sabrina Gordon, Ms. Lucile Padgett, and Ms. Marybeth Thompson, for their help in arranging all academic affairs. Furthermore, my gratitude is also extended to my friends, for their company, support, and encouragement. It is their company that helps to alleviate my feelings of nostalgia. Finally, thank you to all of my family for their support during my study in the United States. Thank you to my parents Dr. Fengxun Lin and Lijuan Yang, for their everlasting love, patience, understanding, support, and encouragement throughout this process. Thank you, Meng Li, for all your support as a wife. I cannot express how lucky I am to have you in my life. Thank you for everything.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ACT	Attention Control Theory
AFOSR	Air Force Office of Scientific Research
AFRL	Air Force Research Laboratory
ALOA	Adaptive Levels of Autonomy
ANOVA	Analysis of variance
CBFV	Cerebral blood flow velocity
CCM	Compensatory Control Model
CI	Confidence interval
CPRS	Complacency-Potential Rating Scale
DSSQ	Dundee Stress State Questionnaire
EEG	Electroencephalography
ERP	Event-related potential
HCT	Human-Computer Trust Scale
Image/IM	Image Analysis task
ISR	Intelligence, Surveillance and Reconnaissance
IST	Institute for Simulation and Training
LOAs	Levels of automation
MAC	Multi-aircraft control
OOTL	Out-of-the-loop
RT	Response time
UAV	Unmanned aerial vehicle

UGV Unmanned ground vehicle

USAF United States Air Force

VIF Variance inflation factor

WR Weapon Release authorization task

OVERVIEW

Automation, Stress, and Trust in UAVs: Overview

Unmanned aerial vehicles (UAVs) have been researched and employed by the United States military services since World War I (Gertler, 2012). The role of UAVs has been growing at an unprecedented rate in the military. UAV missions eliminate the threat to pilots' lives (Gertler, 2012; Stulberg, 2007), and augment combat and surveillance capabilities (Chappelle, McDonald, & King, 2010). Currently, UAVs are serving vital roles in intelligence, surveillance, reconnaissance (ISR) missions and precision strike operations (Chappelle et al., 2010). These roles could be possibly expanded to various "dull, dirty, and dangerous" missions such as air interdiction and aeromedical evacuation (Deptula & Mathewson, 2009). As UAV technology develops, the ability of human operators to manage increasingly automated and sophisticated systems is paramount. This study aimed to contribute to understanding the factors that may determine success or failure in future UAV operations.

Human Factors Issues in UAV Automation

The development of UAVs brings numerous benefits, but it also introduces many human factors issues. Currently, three to four operators are needed in controlling a single UAV. As computers have become more sophisticated, the United States Air Force (USAF) is increasingly interested in automating missions and expects that single operators will be able to manage multiple UAVs with support from automation aids. Working with autonomous systems would face various human factors challenges. Multi-aircraft control (MAC) by a single operator is anticipated to be a particularly time-critical, and cognitively demanding, form of multi-tasking work (Calhoun, Ruff, Draper, & Wright, 2011; Guznov, Matthews, Funke, & Dukes, 2011). In order to understand and maximize the benefits of automation, such as improving task effectiveness and operator performance, an appropriate level of trust in automation must be established and maintained (Lee & See, 2004). Research is needed to better understand how reliance on automation is influenced by potential task design factors, how operator performance interacts with autonomous systems, and which individual difference factors are associated with UAV operator performance.

Modern technology offers automation which promises to increase operator efficiency, enhance the flexibility of operations, and lower workload (Cummings, Brzezinski, & Lee, 2007). However, these benefits require an appropriate level of reliance on automation by operators (Parasuraman & Manzey, 2010). Both over- or under-reliance on automation may compromise the benefits. Empirically, UAV operators show a tendency towards over-reliance on automation technologies, leading to complacency effects in a simulation study (Calhoun et al., 2011). On the contrary, if operators suspect the reliability or functioning of autonomous systems too much, under-reliance may result, limiting the potential benefit and possibly leading to a concomitant increase in operator workload. In prolonged UAV missions, high levels of automation may also induce loss of situational or system awareness by operators (Endsley & Kiris, 1995; Parasuraman, Molloy, & Singh, 1993). This can result in delays or errors when intervention is needed from operators (Wickens & Hollands, 2000).

Additionally, UAV operation may involve considerable workload variation. On the one hand, operators may fail to maintain vigilance due to the inactivity characteristic of many UAV missions, as associated with low task load and lack of interaction with the system (Hancock, Desmond, & Matthews, 2012). On the other, when workload increases, operators are required to allocate their attention among multiple tasks effectively. Generally, automation tends to shift operators from autonomous controllers of work activities to passive monitors of technologies

(Warm, Parasuraman, & Matthews, 2008). Such tasks may elicit passive fatigue on operators, implying a risk of task disengagement that may be exacerbated by fatigue. In some circumstances, high-workload UAV missions may produce active fatigue, which may induce a greater state of distress on operators. Although UAV operation may be exempt from some of the major stressors that afflict traditional pilots, such as fear of physical injury, it may be more psychologically intense and fatiguing.

Individual Differences in UAV Operator Performance

Individual difference factors, such as acquired skills, personality traits, and gender, may influence reliance on automation, fatigue and stress response. Recent research (Spence & Feng, 2010) indicates that video game experience is positively associated with a range of relevant sensory, perceptual, and attentional abilities. Experienced video gamers are found to collaborate with automation more effectively than non-gamers in a simulation environment (Cummings, Clare, & Hart, 2010) In another UAV simulation study, experienced video gamers also showed greater visuospatial attention skills, which may be transferred to the novel environment to improve UAV operator performance (McKinley, McIntire, & Funke, 2011).

Another factor associated with individual differences is personality traits, which may correlate with basic information processing competencies. In a similar domain, all five traits in terms of the Five Factor Personality Model were associated with at least one measure of workload and stress in a simulation of Unmanned Ground Vehicle (UGV) operation (Szalma & Taylor, 2011). In the UAV domain, individuals may interact with automation distinctively. For instance, three groups are categorized as consenters, dissenters, and mixed consenters (Cummings et al., 2010). Generally, consenters tend to follow automation’s suggestion, whereas dissenters usually ignore the automation. Higher degrees of consent are associated with better performance and video game experience (Cummings et al., 2010).

A third relevant factor is gender. The preponderance of male pilots of manned aircraft in the Air Force may reflect both cultural factors and higher aptitude in men, especially for spatially demanding task components (Carretta, 1997; Halpern, 2013). Women are also stereotypically perceived as less resilient. However, how gender differences influence the response to stressors in UAV operation is still unknown. It is also important to disentangle gender differences and video gaming experience since men are more likely to self-identify as serious gamers (Terlecki et al., 2011).

Overview of Study Aims

This study investigated UAV operator performance under two different levels of task demand with the aid of automation at two different levels of automation (LOAs) in a simulation environment. LOA refers to the tradeoff between operator control and delegation of control to the machine. The ALOA (Adaptive Levels of Autonomy; version 3) multi-UAV automation research test bed developed by OR Concepts Applied (Johnson, Leen, & Goldberg, 2007) was used in this study. This desktop-based simulation provided multi-UAV missions, which met the USAF future goal of a single operator managing multiple UAVs, with needed complexity and realism. The task demand was configured by manipulating demands of several secondary tasks. The ALOA test bed also permitted the experimenter to manipulate LOA for specific tasks so that the operator can work with the specific automation aids at different LOAs. Two surveillance tasks (Image Analysis and Weapon Release Authorization) were offered as primary tasks for obtaining performance measurements.

A major finding from automation research indicates that although automation has often improved work efficiency and reduced the burden of work on humans, it is not the case that having more automation (i.e., a higher LOA) is always better (Parasuraman & Riley, 1997). In the UAV domain, how to best apply advanced automation technology to UAV operation remains obscure. Operators are expected to maximize performance, and also minimize any negative consequences of using automation. This study investigated the impact of automation and fatigue on UAV operator performance in a large sample of college students with no prior knowledge of UAV operation. Specifically, this effort looked at the impact of automation and task demand configurations on reliance, trust, and sustained performance, the effect of fatigue on operator reliance on automation, as well as the role of individual differences in reliance on automation and fatigue and stress responses.

INTRODUCTION

Trust and Automation

Automation is the mechanical or electrical accomplishment of work, which replaces functions that are originally performed by humans (Wickens & Hollands, 2000). This replacement could be full or partial, suggesting that automation is not all or none, but can vary across a continuum of levels (Parasuraman, Sheridan, & Wickens, 2000). Levels of automation (LOA), which refer to the tradeoff between operator control and delegation of control to the machine, have been originally identified by Sheridan and Verplank (1978) and adapted and elaborated more recently (Miller & Parasuraman, 2003; Parasuraman et al., 2000). Table 1 shows the LOA model by Parasuraman et al. (2000). Automation, therefore, could vary from offering suggestions, to making decisions, and to action execution. Higher LOA could reduce human workload, but may also cause vigilance decrements, loss of situation awareness, and complacency (Miller & Parasuraman, 2007).

Table 1
Levels of automation model by Parasuraman et al. (2000)

Level	Description
10	The computer decides everything, acts autonomously, ignoring the human
9	Informs the human only if it, the computer decides to
8	Informs the human only if asked, or
7	Executes automatically, then necessarily informs the human, and
6	Allows the human a restricted time to veto before automatic execution, or
5	Executes that suggestion if the human approves, or
4	Suggests one alternative
3	Narrows the selection down to a few, or
2	The computer offers a complete set of decision/action alternatives, or
1	The computer offers no assistance: human must take all decisions and actions

Note. Level 1 is the lowest LOA, level 10 is the highest LOA.

Automation problems are largely due to people's inappropriate level of reliance on automation. Trust plays a vital role on reliance. Many researchers have stated that trust is a mediator between reliability of automation and reliance on automation (Lee & See, 2004; Lee & Moray, 1992; Parasuraman & Wickens, 2008). Generally, higher automation reliability would induce greater trust in automation, which may lead to greater reliance on automation. Although trust has been identified as a belief, attitude, intention, or behavior, in this context, trust is an attitude and reliance is a behavior. Lee and See (2004) defines trust as the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability.

On the one hand, people may trust automation when they should not to, which refers to overtrust, or complacency. Complacency may not lead to a problem until automation malfunctions. On the other hand, people may fail to put sufficient trust in automation when they should, which refers to undertrust, or distrust. Distrust of automation may be due to its complexity or its true low reliability.

Parasuraman and Riley (1997) describe these phenomena in terms of misuse and disuse of automation. Misuse refers to overreliance on automation, which can result in failures of monitoring or decision biases. Disuse refers to the neglect or underutilization of automation, which is commonly caused by false alarm issues. Parasuraman and Riley (1997) also define a third circumstance, automation abuse, which can promote misuse and disuse of automation by human operators. Automation abuse refers to design or management of automation that ignores the consequences for human and system performance and operator's authority.

To describe the relationship between trust and reliance, Lee and See (2004) have distinguished overtrust and distrust in terms of calibration, which refers to how well an individual's trust matches true capabilities of an automation or its trustworthiness. Both over- and distrust are results of poor calibration. Overtrust happens when trust exceeds automation capabilities, whereas distrust results in less trust in automation than its capabilities.

Parasuraman and Manzey (2010) have proposed an integrated model of complacency and automation bias to represent different manifestations of similar automation-induced phenomena, in which attention plays an important role (Figure 1). Complacency potential, which refers to the tendency of a less attentive manner in using automation, is influenced by automation properties (e.g. LOA, reliability) and individual difference factors (e.g. personality traits, attitudes toward technology) (Parasuraman & Manzey, 2010). Furthermore, task context (e.g. workload), individual state (e.g. fatigue state), as well as system properties, may influence attentional bias in using automation due to high complacency potential (Parasuraman & Manzey, 2010).

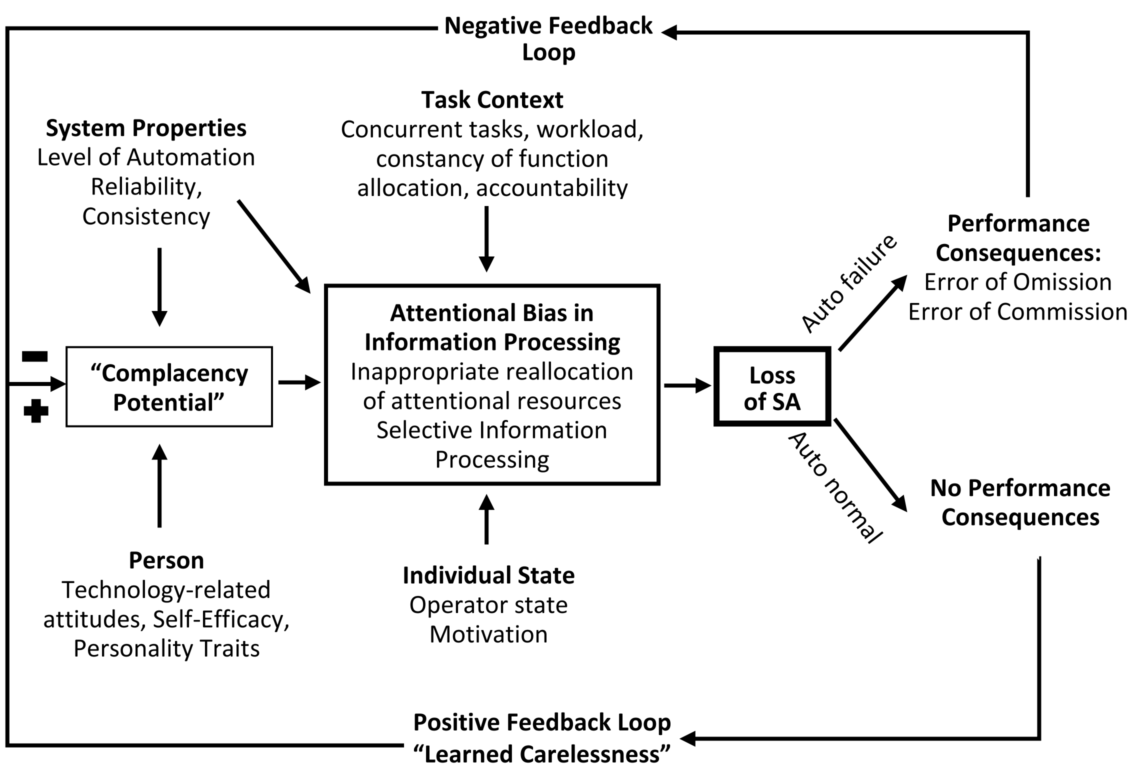

Figure 1. An integrated model of complacency and automation bias (Parasuraman & Manzey, 2010)

The impact of trust in automation reliance may be also affected by other factors, such as individual differences and workload (Hake & Schmid, 1981; Scott, 1980). Self-confidence may be a moderator in the influence of trust in reliance (Lee & Moray, 1992, 1994). Individuals with the perception of their ability beyond their trust in automation's performance would rely on automation less and use more manual control. By contrast individuals with low self-confidence on their ability tend to rely on automation more.

Applications to UAVs/Unmanned Vehicles

Some early UAVs were no more sophisticated than simple radio controlled aircraft managed by human pilots on the ground. In order to achieve the goal of a single operator managing multiple UAVs, automation technologies need to be applied in UAV development. Automated decision support tools, such as decision aids at multiple levels, are critical in facilitating operators in performance and situation awareness (Cummings et al., 2007). Situation awareness refers to 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, 1995). Such automated decision support tools could improve operators' situation awareness by enabling real-time decision-making without continuous human intervention (Hanson & Harper, 2000). Automated decision support tools can be applied to both low and high levels of decision-making tasks, such as target recognition and route planning (Cummings et al., 2007; Drury & Scott, 2008). Decision aiding technology incorporating LOAs may help to reduce operator cognitive load (Parasuraman et al., 2000). Typically, LOAs vary from full manual control to full automation control with intermediate levels, such as management-by-consent, and management-by-exception studied in recent research (Liu, Wasson, & Vincenzi, 2009; Ruff, Narayanan, & Draper, 2002). Management-by-consent, usually, offers a recommendation provided by the automated decision support tool which needs to be either confirmed or changed. Differently, management-by-exception executes the automated decision directly, unless the operator intervenes.

Higher LOAs may enable a single operator to manage more UAVs at the same time, but it also tends to induce out-of-the-loop (OOTL) problems, and leads to poor performance, especially during automation failures (Endsley & Kiris, 1994; Kaber & Endsley, 1997). In addition, higher LOAs might bring vigilance and complacency issues and result in a loss of situation awareness (Endsley, 1996; Endsley & Kiris, 1995; Miller & Parasuraman, 2007). Operators may place excessive trust at higher LOAs and misuse the automation, leading to over-reliance and complacency (Parasuraman & Riley, 1997). An intermediate LOA can lower operator workload and improve performance while helping to maintain situation awareness, supporting consistent performance even as system complexity increases and automation fails (Kaber & Endsley, 1999; Parasuraman & Wickens, 2008; Rouse & Rouse, 1983).

As automation becomes more sophisticated, errors in automation get more difficult to detect, and humans' trust may, consequently, decrease and create undertrust or distrust, leading to disuse of automation (Lee & Moray, 1992; Parasuraman & Riley, 1997; Riley, 1994). Wickens (2000) categorizes unreliable automation into three types: catastrophic, imperfect without awareness, and imperfect with awareness. UAV operators are often aware of the imperfection of the automation. Human response to such imperfect automation depends on the human's allocation of attention, usually visual attention between automation aid and raw information (Moray, Inagaki, & Itoh, 2000; Wickens, 2000). Similar findings indicate that imperfect reliability should not lead to the discarding of automation, but an attention balance strategy between the automation and other relevant information (Merlo, Wickens, & Yeh, 1999; Wickens, 2000; Wickens, Gempler, & Morphew, 2000; Yeh & Wickens, 2000).

Measurement of Trust in Automation

Trust, which originally was used to describe interpersonal activities, is important to be understood and measured since trust may mediate the relationship between individuals and automation just like it mediates relationships between individuals (Sheridan & Hennessy, 1984). Trust in automation can be measured both subjectively, and objectively.

Subjective measures. Although trust and reliance have been identified as two components of attitudes to automation (Singh, Molloy, & Parasuraman, 1993), trust in automation, more generally, may not result in reliance behavior (Parasuraman & Riley, 1997). The combination of attitudes to automation, complacency potential, and particular contextual factors, such as fatigue, high workload, and unfamiliarity of the system, may lead to complacent behavior (Singh et al., 1993). There is no existing scale to measure complacent behavior directly, possibly, due to the difficulty in measuring the behavior subjectively. However, the potential for complacency could be evaluated by attitude ratings towards everyday automation technology (Singh et al., 1993). Singh et al. (1993) have developed a multi-dimensional scale to assess complacency potential, the Complacency-Potential Rating Scale (CPRS). This dispositional scale reveals five factors related to complacency potential, including general attitude toward automation, confidence in automation, reliance on automation, trust in automation and safety in using automation. This study will use the CPRS to understand the impact of individual differences in complacency on human performance across different LOAs and levels of automation reliability in UAV operation.

Subjective situational measures can also be used to assess trust as a consequence of interacting with specific automation. Situational measures of rating trust on specific components or systems are used in a few studies (Bailey & Scerbo, 2007; Lee & Moray, 1992). Jian, Bisantz, and Drury (2000) have identified 12 potential factors of trust between people and automated systems using cluster analysis. They proposed a scale, the Checklist for Trust between People and

Automation, to measure trust in human-machine systems. Additionally, Madsen and Gregor (2000) have found that affect-based trust, including faith and personal attachment, predicts trust well, and developed a psychometric instrument - the Human-Computer Trust Scale (HCT) - to measure human-computer trust. The HCT is designed to measure dispositional trust and is adapted to measure situational trust in UAV missions in this study.

General favorable or unfavorable reactions towards automation do not necessarily predict the actual usage of specific automation systems. Some studies have found that there is no relationship between attitudes to automation and reliance behavior in performance (Singh et al., 1993). Therefore, it is difficult to assess trust in automation only via subjective measures. Objective measures based on operator performance and psychophysiological metrics are also needed.

Objective measures. Although reliance is not completely determined by trust, it is still somewhat guided by trust (Lee & See, 2004). Therefore, trust can be inferred by objectively measuring human performance in terms of reliance. Dixon, Wickens, and McMcarley (2007) have distinguished reliance and compliance, and define reliance as the operator’s action when the automation diagnoses noise in the world, whereas compliance refers to the operator’s action when automation diagnoses a signal in the world. In a broader definition, reliance could refer to operators’ actual usage of automation. In other words, it represents to what extent an operator agrees with a specific automated system. Therefore, more reliant operators should agree with the automation’s recommendations more in using automated decision support systems. In this study, we took the broader definition of an overall agreement with the recommendation to measure reliance on automation decision aids in UAV operations.

Besides performance measures, trust may also be assessed psychophysiologically. Metrics derived from electroencephalography (EEG) and event-related potentials (ERPs) have been used as indices for adaptive automation (Mikulka, Scerbo, & Freeman, 2002; Pope, Bogart, & Bartolome, 1995; Prinzel, Freeman, Scerbo, Mikulka, & Pope, 2003). Another physiological index that could be used to infer trust in automation is eye gaze behavior since the eye gaze behavior of an individual in a task with automation aid indicates the individual's trust in automation indirectly (Flemisch & Onken, 2000; Parasuraman et al., 1993). It is assumed that frequency and duration of scanning may indirectly interpret trust in an automated system's performance. In a UAV simulation study, operators are found to dwell on the automated tasking area more when working with less reliable automation (Wickens, Dixon, Goh, & Hammer, 2005).

Stress and Fatigue

Theories of Stress

Stress, as a vague and complex concept, may refer to actual external stressors to the person's internal reactions, or to the transactional relationship between stressors and stress response (Matthews, 2001). Those stressors can be direct (e.g. noise) or indirect, such as perceived personal incompetence. Those internal reactions could be detrimental to the performance, but sometimes may also be beneficial. Stress can be explained at three levels (Matthews, Davies, Westerman, & Stammers, 2000). At the neural level, stress may be seen as a set of biological responses to challenging stimuli. At the cognitive level, stress may influence the efficiency of information processing. At the knowledge level, stress may be related to motivations and beliefs about the self that influence task strategy.

UAV operators may suffer from multiple sources of stress, such as long hours, shift work, interface difficulties, inefficiencies in control procedures, and conflict between domestic life or personal demands and military operations (Ouma, Chappelle, & Salinas, 2011). The stress may result from the working environment such as exposure to loud background noise from the cooling systems or individual health and sleep issues. However, primarily, the input stress for working operators derives from task demands (Hancock & Warm, 1989). Therefore, the prolonged task itself may also be a great stressor to UAV operators. A UAV operation working shift can last for several hours, so that such task-induced stress may overload and exert time pressure on operators. In addition, working with advanced technology and automation sometimes can be stressful as well. This effort will focus on the acute stress related to UAV operation which involves managing attentional resources to cope with challenging task demands. Loss of attention may lead to vigilance decrement which can be detrimental to operator performance.

Early psychobiological approaches explained stress in terms of the correlation between physiology and emotion. Centralists assert that both physiological and emotional reactions are expressions of central brain systems. Selye (1976) suggests that the "hypothalamic-pituitary axis" is the key brain system related to some long-term stress reactions. Alternatively, peripheralists argue that subjective emotion results from somatic and muscular responses to specific stimulation. Unlike the emphasis of autonomic arousal based on central brain system in centralist approach, peripheralists focus on the conscious awareness of peripheral bodily changes.

Traditionally, the relationship between stress and performance has been explained using the arousal theory. Arousal, generally, refers to individual overall state or level of activities, such as behavioral states (e.g. wakefulness) and emotional states (e.g. tension). The arousal theory is developed from the Yerkes-Dodson Law (Yerkes & Dodson, 1908). Originally, Yerkes and Dodson (1908) draw an inverted-U curve to demonstrate the relationship between the strength of electric shock (a motivating factor) and the speed of learning. The Yerkes-Dodson Law argues that the relationship between arousal level and performance can be expressed as an inverted-U curve. Moderate levels of arousal are optimal for performance. In addition, the optimal level of arousal for performance is inversely related to task difficulty. In other words, harder tasks may require lower arousal level than normal for better performance. Stress may influence arousal level and in turn influence performance.

However, the Yerkes-Dodson Law has not proved entirely satisfactory. Matthews and Amelang (1993) criticize this theory from four aspects, including psychometric, methodological, conceptual, and empirical. Psychometrically, arousal may not be measured reliably and validly. From the methodological aspect, it is relatively easy to fit typical interaction data into such inverted-U curves (Hockey, 1984); therefore, the theory is difficult to falsify. Another difficulty is to decide whether a stressor is arousing or not (Matthews, 1985). Näätänen (1973) also suggests that some stressors may have a distracting effect, which may impair performance through mechanisms other than arousal. The conceptual status of arousal is also criticized. There may be a variety of independent brain systems influencing individual arousal level. Which specific brain systems could be affected by which particular stressors remains unclear. Empirically, data from a variety of studies suggest that the impairment of performance in extreme arousal situations are often weaker than theory expected (Baddeley, 1983; Johnson, 1982; Matthews & Amelang, 1993).

Contemporary cognitive models of stress tend to reject traditional approaches to emotion as being over-simplistic (Matthews et al., 2000). Symptoms, including emotional disturbance, due to stress should be seen as the outcome of an interaction or transaction between individual and environment which develops over time (Cox & Ferguson, 1991; Lazarus, 1991; Lazarus & Folkman, 1984). Lazarus and Folkman (1984) assume that stress results from an imbalance between individual's demands and resources. According to the transactional model of stress (Lazarus & Folkman, 1984), a stressor is only stressful to the individual when it is appraised as likely to tax or exceed the person's coping skills. The same stimuli may be appraised differently across individuals and contexts. Appraisal includes interpretations of the stressors and analyses of the available resources. Coping skills may involve active efforts to regulate the external situation (task-focused coping) or somewhat less effortful responses such as rethinking one's attitude to the potential stressor (emotion-focused coping) or trying to avoid it totally (avoidance). Therefore, whether an event is stressful or not is not solely a property of external stimuli. In the performance context, a critical issue is whether the person appraises their coping abilities as adequate to maintain a personally-acceptable standard of performance, given prevailing task demands (Matthews, 2001). Lazarus and Folkman (1984) state that performance and stress are dynamically interrelated. Potentially, stress can impair or improve performance, but the Lazarus and Folkman (1984) theory does not provide a detailed account of performance impacts.

Hockey (1997) has proposed another cognitive-energetical framework, the Compensatory Control Model (CCM) of performance under stress, which accounts for the effects of stress on performance. Hockey (1986) argues that 1. performance is often maintained under stress; 2. the stress effects depend on the appraisal of stress and vary for different stressors and task demands; 3. the relationship of stress and activation depends on the level of task engagement. The CCM assumes that performance is goal oriented; goal states are managed by self-regulatory; and regulatory activity is resource consuming. The model contains two feedback loops (see Figure 2). The lower loop A controls performance more or less automatically when only little effort or mental resources are required for the activity. The upper loop B may be engaged when the task becomes demanding. The effort monitor detects demands on regulatory activity. When a discrepancy is detected, the supervisory controller can either shift resources to maintain the task goal or change goals strategically for the task. Stress factors may elicit various changes to system operation that impact performance. For example, some stressors produce "strain" as the person actively compensates for increased processing demands by increasing effort. By contrast, fatiguing agents may lower task goals and lead to effort-reduction.

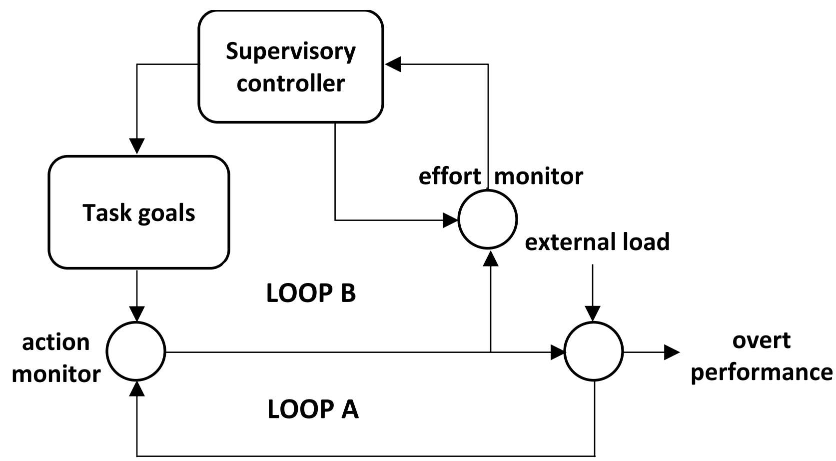

Figure 2. Compensatory Control Model from Hockey (1997)

Workload as a Stress Factor

UAV operations involve considerable workload variation which may lead to stress, and in turn, influence operator performance. Hancock and Warm (1989) developed a theoretical dynamic model for stress and performance. Individuals can adapt effectively to some levels of stress without showing significant performance decrement. However, both extreme overload and underload could result in failures in such adaptation. These adaptations are illustrated as a series of extended inverted-U curves in the model. At the psychological level, the adaptability is related to individual's attentional resource capacity. Stress can result in a reduction of available attentional capacity, especially when the task environment is not configured to support compensatory or coping efforts (Hancock & Warm, 1989). UAV mission tasks vary in the levels of workload demands, including both high and low workload, from time to time. Chronically high workload may contribute to stress, whereas low workload and monotony may induce fatigue. If the stress of operators exceeds their optimal adaptability due to either of these inappropriate work configurations, the result may be catastrophic. From the perspective of the multiple resource theory (Wickens, 1984), UAV operations require attention from multiple resource pools, including visual, auditory, verbal, and spatial. This research focused on visually-demanding surveillance tasks. Operators in UAV missions often need to maintain a high level of vigilance, which requires hard mental work and is stressful (Warm et al., 2008). Operators' vigilance decrement may be primarily controlled by workload (Warm, 1993). Prolonged UAV missions may deplete the pool of attention resources as the operators get stressed and fatigued. Vigilance tasks, such as the detection tasks in UAV missions, may also reduce task engagement and increase distress level, especially in demanding workload scenarios (Miller, 2012).

The relationship between UAV automation, workload and stress is potentially complex. In general, automation should alleviate workload and stress by keeping cognitive demands to a manageable level. Indeed, automation is seen as a key to future UAV operations in which a single operator controls multiple vehicles (Mouloua, Gilson, & Hancock, 2003). However, despite automation support, the multi-UAV operation may still exacerbate the stress induced by workload-related factors (Cummings, Mastracchio, Thornburg, & Mkrtchyan, 2013). Automation may fail to mitigate workload if it is not used appropriately. As discussed next, automation may also increase the operator's vulnerability to fatigue and loss of situation awareness (De Winter, Happee, Martens, & Stanton, 2014).

Cognitive Fatigue

When managing highly automated UAVs, much of the operator's workload derives from passively monitoring mission progression, system status, alert of malfunctions, and other parameters (Mouloua, Gilson, Kring, & Hancock, 2001; Tvaryanas, Thompson, & Constable, 2006).

Although UAV operation contains long periods of low workload (Cummings et al., 2013), it may also require intense activities for brief periods (Cummings et al., 2007). Such workload variation, according to Desmond and Hancock's (2001) theory, may induce different forms of cognitive fatigue.

Desmond and Hancock (2001) distinguished two types of fatigue, active and passive fatigue, associated with different cognitive workload levels. Specifically, active fatigue refers to the state change resulting from "continuous and prolonged, task-related psychomotor adjustment", whereas passive fatigue develops when performing system monitoring with either rare or even no overt perceptual motor requirements (Desmond & Hancock, 2001). The properties of some UAV operations, such as prolonged ISR missions, may trigger such cognitive fatigue.

Different forms of cognitive fatigue may differ in their effects on UAV operators' performance. A recent study (Saxby, Matthews, Warm, Hitchcock, & Neubauer, 2013) suggests that active fatigue is associated with distress, overload, and heightened coping efforts, whereas passive fatigue links to the loss of task engagement, cognitive underload, and reduced challenge appraisal. Passive fatigue may pose greater detrimental effects on performance than active fatigue. For instance, the recent simulated driving study (Saxby et al., 2013) reveals that drivers under passive fatigue show slowed responding, such as delayed brake and longer steering reaction time, to emergency events, whereas active fatigue has a little performance impact. Passive fatigue may be more harmful to the individual's alertness due to the loss of attentional resources (Warm et al., 2008) or strategic reduction in the allocation of effort (Hockey, 1997).

Trust and Fatigue

The impact of fatigue on trust in automation has been neglected in prior research and remains unclear. Generally, automation is designed to be supportive to UAV operators, especially in stressful and fatiguing circumstances. Potentially, automation might alleviate stress by reducing cognitive load. Conversely, passive fatigue might be relieved if the automation is able to handle monotonous task requirements, such as maintaining vigilance for rare events. Thus, fatigue does not necessarily impact trust adversely, but some concerns remain.

One hypothesis is that operators under passive fatigue may show excessive trust on automation. Hockey's (1997) CCM model, described previously, links fatigue to reduced performance standards and a reduction in proactive effort to maintain standards. These processes may lead to increased reliance on automation as the person reduces effort directed towards maximizing performance. Consistent with this hypothesis, fatigued drivers are more likely to use optional automation than non-fatigued, even though it does not enhance performance, in a simulated surface vehicle study (Neubauer, Matthews, Saxby, & Langheim, 2011). Probably, such over-reliance on automation under fatigue is especially pronounced when the automation is highly reliable.

An alternate view derives from the observed impact of automation on passive fatigue and the loss of task engagement (Saxby et al., 2013). The impairment of attention may interfere with operator's ability to monitor and manage automation effectively. In this case, the operator may be vulnerable to under-trust as well as to over-trust of automation. Especially if the automation is perceived as unreliable, fatigued operators may not apply sufficient effort to evaluate it further, so that under-reliance on automation or totally ignoring the automation may occur.

In sum, although fatigue, especially passive fatigue, may encourage over-reliance on automation, in some other instances, fatigue might also lead to neglect of automation. This effort will examine the effect of fatigue on operator reliance on automation in UAV domain in a simulated environment to provide further evidence on this issue.

Individual Differences in Stress

A major challenge to understanding the impact of stress and fatigue on the UAV operator is that individuals differ considerably in their responses to complex task environments (Szalma, 2009). Relevant individual difference factors include both stable traits that define personality and transient subjective states of stress and fatigue. Gender and task-relevant skills are also potential sources of variability. Stress is sometimes considered as a unitary construct: for example, the personality trait of neuroticism is associated with a general vulnerability to situational stress response (Matthews, Deary, & Whiteman, 2009). However, in the human factors context, it is often productive to discriminate different components of stress and fatigue that may be differently related to performance outcomes (Matthews, 2016). This section reviews some of the multiple individual difference factors that may be relevant to the UAV operator.

Three-Factor Model (DSSQ)

The Dundee Stress State Questionnaire (Matthews et al., 2002) is developed for investigating task-induced stress based on a three-factor model raised by Matthews and colleagues (Matthews, 2016; Matthews, Joyner, Gilliland, et al., 1999; Matthews et al., 2002). Factor analysis reveals a two-level model. First-level factors distinguished 11 dimensions of subjective states. That is, there are a variety of ways in which "stress" may be experienced. By using factor analysis of state scales to group the inter-correlated first-level or primary factors, three second-level factors are integrated across three different domains, including motivation, cognition, and affect. The three-factor model suggests that task stress may be experienced in three different transient states, labeled as task engagement, worry, and distress. Task engagement represents energy, motivation, and alertness, whereas low task engagement indicates tiredness, loss of interest in the task, and distractibility. Worry, as a cognitive factor, corresponds to self-focused attention, low self-esteem, and high cognitive interference. Distress refers to high tension, unpleasant mood, and low confidence and perceived control.

Stressful tasks can induce a variety of subjective state responses, such as increases in distress, increases in worry, and decreases in task engagement (Matthews, Szalma, Panganiban, Neubauer, & Warm, 2013). The multidimensional pattern of response varies according to task demands (Matthews, 2016). The UAV operation features considerable workload variation. The operator may monitor the system under conditions of low workload and monotony for a long period, whereas high cognitive workload is imposed immediately when a target is detected or an emergency is declared. A large number of studies (Langner, Steinborn, Chatterjee, Sturm, & Willmes, 2010; Matthews, Warm, Reinerman-Jones, et al., 2010; Matthews & Campbell, 2010; Teo & Szalma, 2011; Warm et al., 2008) suggest that high workload tasks, even of short duration, can lead to increases in distress easily (Matthews et al., 2013). Although workload factors, such as multitasking in UAV operation, can elevate distress, distress may not be driven directly by workload. For example, lower maneuverability in UAV simulated control elevates operator's workload and impairs task performance, but has no effect on distress (Guznov et al., 2011). Similarly, Szalma et al. (2006) observed increased distress after a stressful vigilance task, but knowledge of results format in feedback had no impact on distress. That is, it may be the appraisal of the manageability of demands, rather than the objective level of demands that drives stress response.

Generally, task engagement reflects effort committed to achieving task goals (Matthews et al., 2002). In the view of cognitive resource theory, task engagement may relate to the availability of a general attentional resource. In a vigilance study, the evidence of convergence between performance, task engagement, and psychophysiological indices, especially cerebral blood flow velocity (CBFV), supports resource theory (Matthews, Warm, Reinerman-Jones, et al., 2010). Declines in task engagement can occur in both short-duration vigilance tasks and prolonged monotonous tasks. In a simulated driving study, for instance, a large-magnitude decline in task engagement is observed after brief and more prolonged periods of automated driving (Saxby et al., 2013). Although stressful tasks usually impair task engagement, challenging tasks or game-like elements in complex tasks may elevate task engagement (Matthews et al., 2013). A good example is that Guznov et al. (2011) found elevated task engagement in a simulated UAV study.

Worry usually declines during general tasks. DSSQ contains four scales for worry factor, including self-focus, self-esteem, task-irrelevant cognitive interference, and task-relevant cognitive interference. Typically, self-focus decreases, self-esteem increases, and task-irrelevant cognitive interference decreases in general tasks (Matthews, Joyner, Gilliland, et al., 1999). Exceptionally, worry tends to be maintained or even elevated in fatiguing driving tasks. For example, no significant change in worry was observed after a monotonous simulated driving task (Neubauer, Matthews, Langheim, & Saxby, 2012). Also, task-irrelevant cognitive interference was elevated among long-haul truck drivers during the approximate 12-hour shift (Desmond & Matthews, 2009). Automation, such as adaptive cruise control, may contribute to increased task-irrelevant cognitive interference score in vehicle driving (Stanton & Young, 2005). In the UAV context, monotonous missions may be associated with the mind-wandering that appears to accompany worry (Cummings et al., 2013).

The DSSQ offers two versions (full version & short version). The short version DSSQ with 21 items measures the three second-level factors only, including task engagement, distress, and worry, while the full version provides additional details on first-level factors. There is evidence supporting the validity of the DSSQ as an assessment instrument on profiling stress response to task performance and profiling individual differences in response to a variety of human factors contexts (Matthews, 2016). As a subjective measure, DSSQ scores are still predictive for performance even when psychophysiological factors are controlled (Abich, Matthews, & Reinerman-Jones, 2015).

Personality and Stress

The Five Factor Model of personality is often used as a basis for the assessment of stable individual differences in stress response. The Five Factor Model contains five factors grouped by factor analysis to describe the individual's personality. These five factors are Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Considering human performance and stress response, most findings are focused on Extraversion and Neuroticism (Eysenck & Eysenck, 1985), but few studies have been done on the other three factors (Matthews, Deary, & Whiteman, 2003).

Extraversion. Extraversion refers to the characteristics of social interaction, such as activity, assertiveness, warmth, gregariousness, and positive emotions. Matthews and colleagues (Matthews et al., 2003) have identified that extraversion has the advantages of greater working memory, divided attention, and resource capacities, but extraverts also tend to be poorer in sustained attention and more lenient in choosing response criteria. Due to this general tendency, extraversion should negatively correlate with workload and stress in missions requiring divided attention to multiple displays or tasks such as UAV operations (Szalma & Taylor, 2011). Also, in terms of stress response, this trait is often related to lower post-task distress level (Matthews, Joyner, Gilliland, et al., 1999).

Neuroticism. Neuroticism refers to the individual's tendency to experience unpleasant or negative emotions, such as anger, anxiety, depression, and sadness. Typically, this trait is associated with greater vulnerability to stress, such as higher distress and worry (Matthews, Joyner, Gilliland, et al., 1999). In terms of dealing with stressful tasks, Matthews and Campbell (1998) have found that neuroticism is correlated with emotion-focused and avoidance coping style. The complexity of UAV task components require appropriate working memory, and attentional resources, but individuals high in neuroticism tend to be more vulnerable to impairment of working memory, attentional resources, and sustained attention (Matthews et al., 2003).

Conscientiousness. Conscientiousness refers to the individual's tendency to be organized and dependable. Conscientious individuals usually show self-discipline, act dutifully, and aim for achievement. Generally, this trait is positively related to performance, and individuals high in conscientiousness perform better to achieve goals and perceive lower levels of stress and workload when the environment supports the task goal (Szalma & Taylor, 2011). In terms of the stress response, conscientiousness often predicts greater task engagement and lower distress and worry (Matthews, Joyner, Gilliland, et al., 1999). Automation reliability may moderate the effect of conscientiousness on performance and stress response. When the automation aid is reliable, conscientiousness should predict better performance, and conscientious operators should be less vulnerable to complacency, and misuse or disuse of automation (Szalma & Taylor, 2011).

Agreeableness. Agreeableness is a tendency to be compassionate and cooperative toward others. Individuals high in agreeableness usually perform better in tasks requiring interpersonal interaction and cooperation (Szalma & Taylor, 2011). In dealing with potentially stressful task demands, high agreeableness individuals often use less avoidance coping strategies (Matthews &

Campbell, 1998). Agreeable individuals also tend to experience lower distress (Matthews, Joyner, Gilliland, et al., 1999).

Openness. Openness reflects individuals' degree of intellectual curiosity, creativity, and preference for novelty. Generally, openness predicts better performance and lower perceived workload and stress, especially in tasks with novel situations or environments (Szalma & Taylor, 2011). In terms of the stress response, openness usually is negatively associated with distress (Matthews, Joyner, Gilliland, et al., 1999). Automation properties, such as reliability, may moderate the effects of openness on performance. Individuals high in openness may perceive higher workload and stress in highly reliable automated aided tasks due to insufficient cognitive stimulation, but they may be less vulnerable to misuse of automation (Szalma & Taylor, 2011).

Performance Correlates of Stress States

Stress states can reflect both direct physical stressors, such as noise, and indirect stressors, such as perceptions of task demands and physiological responses. Changes in those states may influence information processing factors, including basic cognitive parameters (e.g., working memory, attentional capacity) and strategic factors (e.g., understanding of the task, strategies to achieve task goals), and in turn influence performance (Matthews et al., 2013).

Task engagement. A Large number of studies have demonstrated that the state of task engagement is predictive for task performance requiring attentional resources. Matthews et al. (2013) have summarized that the task engagement – performance correlation is typically around 0.3. Studies of tasks sharing similar components with UAV operations suggest that task engagement is associated with better vehicle control in a moderately fatiguing simulated driving study (Funke, Matthews, Warm, & Emo, 2007), and predicts perceptual sensitivity in vigilance tasks (Matthews, Warm, Reinerman-Jones, et al., 2010; Matthews, Warm, Shaw, & Finomore, 2010).

In terms of the impact on vigilance, evidence from structural equation modeling (Helton, Matthews, & Warm, 2009) has shown that task engagement mediates the effects of external stressors, such as loud noise, on vigilance factors. Besides task-processing factors, stress states may also influence strategic factors, such as coping strategy. For example, in a simulated driving study (Neubauer et al., 2012), drivers with low task engagement appeared to be more likely to use automated driving voluntarily to reduce task load, which indicates that the fatigued performer may lower task goals (Hockey, 1997).

Distress. Distress is expected to be detrimental to attention (Matthews & Campbell, 2010). This detrimental effect on performance is found in a few vigilance studies (Shaw et al., 2010), although task engagement is a more reliable predictor of vigilance (Matthews et al., 2013). Distress may also impair an individual's working memory and multi-tasking. Evidence has been found in a longitudinal study using the Turner and Engle (1989) task (Matthews & Campbell, 2010). In the unmanned vehicle context, Abich et al. (2015) found that distress was associated with poorer detection performance in task scenarios that required multi-tasking. Besides attention and working memory, distress interferes with executive control as well. Matthews and Zeidner (2012) have confirmed that distress is associated with poorer inhibition of task-irrelevant stimuli. On the contrary, the beneficial effect of distress is also seen is some real life contexts. For instance, distress was reported to be correlated with greater accuracy in a police handgun shooting exercise (Stafford, Oron-Gilad, Szalma, & Hancock, 2004). In Hockey's (1997) model, distress may be associated with compensatory effort as the person attempts to cope with high task demands.

Worry. Results from test anxiety research in the educational context suggests that worry generally impairs attention, working memory, and information retrieval from long-term memory

(Zeidner, 2010). Like the distress factor discussed above, worry also shows inconsistency in some associations. Matthews and colleagues (2012) have identified that worry correlates with perceptual sensitivity only in a cognitive vigilance task, not in a sensory vigilance task. Pre-task worry only reflects the impairment of arithmetic recall in a working memory task but does not predict the performance of verbal recall (Matthews & Campbell, 2010). In terms of executive control, worry may slow performers in switching tasks (Matthews et al., 2013). In a simulated driving study (Funke et al., 2007), worry was predictive of poor vehicle control, which may apply to UAV operation as well.

Gender and Video Gaming

Video gamers may be superior in aptitudes or skills for operating UAVs or other automated systems. Recent studies demonstrate that video game exposure is positively associated with a range of sensory, perceptual, and attentional abilities (Spence & Feng, 2010), which are identified as critical aptitudes for UAV operation (Chappelle et al., 2010). Spence and Feng (2010) also suggested that training on video games improves performance on other spatial tasks unrelated to the training game. This transfer effect is also seen in the UAV domain. For example, in a simulated UAV study, experienced video gamers showed greater visuospatial skills than actual UAV pilots (McKinley et al., 2011). Experienced gamers also show strengths in interacting with automation. Findings from a recent study (Cummings et al., 2010) suggest that video gamers could collaborate more effectively with automation in simulated UAV missions. Spence and Feng (2010) have categorized video games into three types, including action, driving, and maze or puzzle games, based on cognitive demands. Among those, action games may especially share a variety of critical aptitudes for UAV operations, such as speeded information processing, visual perception, and various forms of attention and spatial procession. Therefore, video gaming experience is potentially beneficial to UAV operations.

Traditionally, military pilots are mostly male. This may reflect both cultural factors and higher aptitudes, such as spatial processing, in men (Carretta, 1997; Halpern, 2013). However, considering the differences between traditional piloting and UAV operations, the gender differences in piloting manned vehicles may not generalize to managing unmanned systems. In an occupational study using a real UAV operator sample, no gender differences in emotional exhaustion were found (Chappelle, Salinas, & McDonald, 2011). Gender differences in stress response in UAV operations under various workload levels still need to be examined. Since men are more likely to self-identify as serious gamers (Terlecki et al., 2011), it is important to disentangle gender differences and video gaming experience as well. Findings for individual differences in gender and other factors may help to target potential UAV operators for future recruiting.

Aims of Study

Generally, this study aimed to develop a simulation environment for studying the role of individual differences in UAV operation under different task configurations. Specifically, the study aimed to determine the impact of workload and levels of automation (LOAs) on UAV operator performance, stress response, and operator reliance on automation. It also aimed to examine the role of individual difference factors associated with gender, video gaming experience, personality, and trust in simulated UAV operation, and their dependency on task factors.

Aim 1. Examine the Impact of Levels of Automation (LOAs) on Task Performance and Operator Reliance

Higher LOAs reduce operator workload, but may impair vigilance and situation awareness, and also lead to complacency (Miller & Parasuraman, 2007). In this study, high and low

LOAs with the same relatively high reliability were applied to examine the impact of LOAs on task performance and operator reliance. Specifically, the study contrasted management-by-exception (Level 6 in Parasuraman et al.'s LOA model) with management-by-consent (Level 4 in Parasuraman et al.'s LOA model). It was hypothesized that operators should show higher reliance on automation and better performance when using the higher level of automation.

Aim 2. Examine the Impact of Task Demand on Task Performance and Operator Reliance

Higher task demand should elicit higher workload, and in turn induce distress poor performance, whereas low task demand should elicit lower workload and may trigger loss of task engagement in operators (Desmond & Hancock, 2001; Saxby et al., 2013). In this study, the frequency of secondary tasks was varied to manipulate task demand. It was hypothesized that high task demand should have detrimental effects on performance, and operators under low task demand should show more reliance on automation.

Aim 3. Examine the Role of Individual Differences in Simulated UAV Operation

Previous research has suggested that individual differences may have impacts on operator response in terms of acute stress, performance, and reliance on automation. For example, video gamers are found to be more collaborative with automation in a simulated UAV task (Cummings et al., 2010). In a simulated ground vehicle task, all five personality traits show correlations with at least one measure of perceived workload and stress (Szalma & Taylor, 2011). Transient engagement and distress are associated with performance in a UGV simulation (Abich et al., 2015). The aim of this study was to investigate relationships between video gaming experience, personality, gender, trust, performance, subjective stress response, and reliance on automation. It was hypothesized that gaming experience should correlate with performance and lower levels of fatigue; personality and stress states should predict task performance and reliance on automation; task performance and reliance on automation should differ between men and women.

Aim 4. Examine Moderators of Individual Differences

Associations between individual difference factors and performance during unmanned vehicle operations may vary in different task configurations (Szalma & Taylor, 2011). Specifically, skills associated with video gaming, as well as adaptive stress states, may be most advantageous under high task demand circumstances. Thus, the study aimed to test whether task demand moderates the associations between individual difference factors and performance. It was hypothesized that task demand should moderate the associations between individual difference factors and performance. Individual difference factors, such as gaming experience and personality, may be more predictive under high task demand. The moderator effect of LOA was investigated on a more exploratory basis.

Aim 5. Examine the Correlates of Subjective Trust

Automation with high reliability is designed to reduce workload, alleviate stress, and optimize operator performance, but it may result in complacency and situation awareness problems (Miller & Parasuraman, 2007). Individual differences may have an impact on operator interacting with automated systems. For example, gaming experience and personality factors have been shown to influence performance in tasks with automation (Cummings et al., 2010; Szalma & Taylor, 2011). This study examined the possible correlates of subjective trust, such as gaming experience, personality, and performance metrics. It was hypothesized that subjective trust should correlate with reliance on automation.

METHODS

Study Design

A 2 (task demand: high versus low) × 2 (LOA: management-by-consent versus management-by-exception) between-subjects factorial design was adopted in this study.

Participant Recruitment

A total of 101 participants (59 women, 42 men, M_{age} = 18.95, *SD* = 1.80) were recruited from the University of Central Florida undergraduate psychology student pool via the SONA system. Student participants received course credits for participation. Participants were healthy individuals between 18 and 40 years old representing the age group and educational level of the enlisted military service core that may be recruited for future UAV operations. Participants who may be vulnerable to adverse reactions, such as excessive stress, resulting from the test environment were excluded. All participants reported having normal or corrected to normal vision, color vision, normal hearing, and English fluency.

Lab Space and Equipment

A desktop workstation was utilized for this study. The UAV simulation was run on a custom-built desktop with 4th generation Intel® Core™ i7 CPU, dual 24-inch LED-backlit widescreens (1920 × 1200 resolution), two stereo speakers, and standard mouse and keyboard.

UAV Simulation

The ALOA (Adaptive Levels of Autonomy) multi-UAV research test bed developed by OR Concepts Applied (Calhoun et al., 2011; Johnson et al., 2007) was used for the study. This simulation supports task manipulations representing UAV operations in needed complexity and realism. Nine tasks (Table 2) were designed to represent the task demands for a single operator managing four UAVs with an automation aid at the same time. The LOAs were varied in two intermediate levels with high reliability (correct 80% of the time) for the primary tasks. Management-by-consent required participants to accept or change the option recommended by the automation. Alternatively, with management-by-exception, the system was set to act on the option recommended by the automation automatically unless a different option was selected before the availability of operator response was timed out (30 or 20 seconds based on tasks).

Table 2
Task priorities, actions, LOAs, and measures

Task Type	Priority	Operator Action	LOA	Measures
Target Allocation	1	As new imaging tasks are added, allocate the new tasks within the existing tasks/UAVs	Manual	RT/accuracy
UAV Rerouting	1	Based on Allocation, select, confirm, acknowledge, or initiate new plans based on current rules of engagement	Management-by-consent	RT/accuracy
Image Analysis	2	Identify number of targets (and click to confirm)	Management-by-consent / Management -by-exception	RT/accuracy
Weapon Release Authorization	2	Identify if target is present or absent (and click for authorization or not)	Management-by-consent / Management -by-exception	RT/accuracy
Unidentified Aircraft	3	Click red plane symbol when presented	Manual	RT/accuracy
Compare Digit Pairs	4	Determine whether the digits meet certain criteria and response	Manual	RT/accuracy
Respond: Audio Chatter	4	Respond color number combination if certain call sigh is prompted	Manual	RT/accuracy
Respond: Visual Status	4	Click on yellow or red colored light for health status	Manual	RT/accuracy
Retrieve Information	4	Answer questions in chat window using vehicle status information	Manual	Accuracy

Management-by-consent: accept or change the option recommended by automation.
Management-by-exception: the system automatically acts the option recommended by automation unless a different option is selected before timed out.
RT: response time.

Table 3

Tasks manipulated across low and high task demand conditions

	Task Demand	
Task and Frequency in Trial	Low	High
Retrieve Information	10	80
Respond: Visual Status	30	240
Respond: Audio Stream	32	240
Compare Digit Pairs	10	80
Monitor Chat Noise	20	180

Note. Numbers refer to the number of tasks in one-hour session of experimental trial

Task frequencies of the secondary tasks (Table 3) were manipulated to create task demand variation across conditions in one hour long experimental scenarios whereas task frequencies of two primary surveillance tasks were held constant. There were 6 tasks or 14 tasks per minute to induce low and high task demand respectively. Most secondary tasks required responses to visual or audio signals, searching and retrieving information, or comparing digit pairs. All tasks were displayed in the certain panel of the simulation window (Figure 3). Primary surveillance tasks were signaled by adding a taskbar with a timer showing time remaining in the task window. Image Analysis and Weapon Release authorization tasks were timed for 30 and 20 seconds respectively. The taskbar would be blanked and the task response would be recorded as a "miss" if there is no operator response before task availability was timed out.

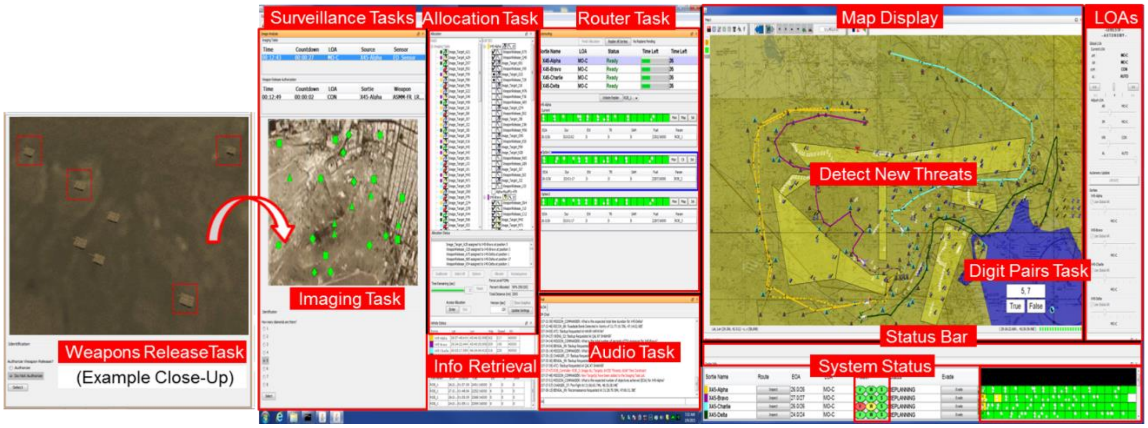

Figure 3. Task interface for multi-UAV operation in the ALOA simulation

In the Image Analysis task, images taken by a UAV with an overlay of 19-26 green symbols varying in shapes, including diamonds, squares, circles, and triangles, were shown in the task panel. Participants were asked to identify the number of diamonds and select the number from eight options. The automated aid system recommended one from the eight options by highlighting it. The reliability of the automation was set to be 80% correct.

In the Weapon Release authorization task, participants were asked to distinguish hostile tanks from allied tanks and detect whether the hostile tanks in given picture were correctly marked. The tanks differed in body width and barrel length subtly. The pictures were degraded in quality to increase the difficulty in discrimination. The automation aid system recommended one option from "authorize" or "do not authorize". Also, reliability was set to be 80% correct.

Subjective Measures

Demographics Questionnaire (APPENDIX A)

The demographics questionnaire contains 21 items. The questions ask about a range of biographical information, including age, gender, health status, education level, computer usage and expertise, and video gaming experience and expertise.

40 Mini-Marker Personality Scale (APPENDIX B)

The 40 Mini-Marker Personality Scale measures personality traits based on the Five Factor Model, in terms of openness, conscientiousness, agreeableness, extraversion, and neuroticism. The scale consists of 40 common human traits. Participants were asked to rate how accurately these 40 traits described themselves in general using a 9-point Likert scale ranging from "0 = Extremely Inaccurate" to "9 = Extremely Accurate". This 40-item scale is a brief version of Goldberg's (1992) 100 adjective markers for personality. Compared to the original scale, it has less difficult items, lower inter-scale correlations, with no loss of validity.

Complacency Potential Rating Scale (CPRS; APPENDIX C)

The Complacency Potential Rating Scale (Singh et al., 1993) is a multi-dimensional scale for assessing the individual's dispositional propensity to grow complacent in using automation. This 20-item scale measures four components of complacency, including confidence-related, reliance-related, trust-related, and safety-related complacency. Every item has a statement about an attitude toward common systems with automation technology (e.g. "Even though the automatic cruise control in my car is set at a speed below the speed limit, I worry when I pass a police radar speed trap in case the automatic control is not working properly"). Participants were asked to indicate how much they agreed with each statement using a 5-point Likert scale ranging from "0 = Extremely disagree" to "4 = Extremely agree".

Dundee Stress State Questionnaire (DSSQ: short version; APPENDIX D)

The short version of the Dundee Stress State Questionnaire (DSSQ) measures three higher order dimensions of subjective states in terms of task engagement, distress, and worry. In this study, it was administered to gauge the stress response elicited by task load manipulation. This questionnaire was administered both before the task as a baseline measure, and after the task reflecting the state in the final 10 minutes of experimental task. The DSSQ contains 30 items about feelings and thoughts. Participants were instructed to rate how accurately those statements described their current emotional states using a 5-point Likert scale ranging from "0 = Definitely false" to "4 = Definitely true".

Metrics for Trust in Automation (APPENDIX E)

The Metrics for Trust in Automation is a 22-item survey developed by the Air Force Research Laboratory (AFRL) for studies using the ALOA simulation. The first seven items addressed the general feedback on the simulated UAV operation in aspects of task difficulty, confidence in performance, trust in automation, workload, and training adequacy. The following 15 items focused on the automated aid in three primary tasks, including rerouting and two surveillance tasks. Questions covered competence of the automation, accuracy of the automation, trust on the automation, consistency of the automation, and confidence for the automation. Five-point Likert scales (descriptions varied by questions) were used for answering the questions.

Human - Computer Trust Scale (APPENDIX F)

The Human - Computer Trust Scale for this study was adapted from the Human - Computer Trust Scale (Madsen & Gregor, 2000). This 9-item scale measures trust in automation from affective and cognitive aspects. Participants were asked to evaluate their perceived reliability, perceived technical competence, perceived understandability, faith and personal attachment in automation, as well as global trust in automation using a 5-point Likert rating scale ranging from "0 = Extremely disagree" to "4 = Extremely agree".

NASA - Task Load Index (NASA-TLX; APPENDIX G)

The NASA - Task Load Index (Hart & Staveland, 1988) is a widely used multi-dimensional measurement of subjective workload. It consists of six rating scales for workload-relevant factors, including mental demand, physical demand, temporal demand, performance, effort, and frustration. All factors, except performance, are rated on a 0 - 100 scale from "Low" to "High". Performance is rated on a 0 - 100 scale from "Good" to "Poor".

Procedure

Pre-Task Activities

Before the experiment sessions, an informed consent agreement was received by researchers. Then, participants were asked to turn off cell phones and remove watches. Next, participants were instructed to complete the pre-task survey set, including the Demographic Questionnaire, the 40 Mini-Marker Personality Scale, the Complacency Potential Scale, and the pre-task DSSQ. The total time for pre-task activities was approximately 20 - 30 minutes.

Training

After completing pre-task surveys, training started with an introduction using PowerPoint slides, followed by a live simulation demonstration and hands-on practice. In the training slides, the interface of the simulation, task priority, and every task operation in the simulation were briefly illustrated. In the live simulation demonstration, every function of control and task was explained in detail. Finally, participants needed to practice with the live simulation under supervision. They had a “cheat sheet” about all the tasks for quick reference and were able to ask any questions during the training. Researchers monitored the practice process to ensure that participants understood all the tasks and were qualified for the experimental task. A second hands-on practice could be run if needed. But this was never performed. Participants were allowed to take a break after the training session. Training took approximately 60 minutes.

Experimental Task

Participants were randomly assigned to one of the four conditions. Before the experimental task, researcher repeated instructions for simulation controls briefly and emphasized task priorities. Participants were not allowed to interact with researchers during the 1-hour experimental task. Researchers confirmed with participants that nothing remained unclear before proceeding to the experimental task. The experimental task ran for 60 minutes.

Post-Task Activities

After the experimental task, participants were instructed to complete the post-task survey set immediately. Post-task survey set consisted of the post-task DSSQ, the Metrics for Trust in Automation, the Human - Computer Trust Scale, and the NASA - Task Load Index. Finally, before dismissing participants, researchers answered any concerns, asked for verbal feedback, and provided the research study evaluation survey from the psychology department. Post-task activities took approximately 15 minutes. All the sessions in total were completed within three hours.

RESULTS

The Impact of LOAs and Task Demand on Subjective States

Workload

Bonferroni-corrected *t*-tests were run to test the effects of experimental manipulations. It was confirmed that workload (NASA-TLX global workload) was significantly higher in high task demand conditions (M = 57.1) than in low task demand conditions (M = 46.2), t (99) = -3.52, p = .001. According to NASA-TLX, the manipulation of task demand successfully elicited higher workload in all aspects, including mental demand, t (99) = -1.78, p = .079; physical demand, t (75.9) = -3.77, p < .01; temporal demand, t (99) = -2.43, p < .05; effort, t (99) = -2.47, p < .05, and frustration, t (99) = -2.73, p < .01, in high task demand conditions (Figure 4). However, there was no difference in self-reported performance, t (99) = -.21, p = .835.

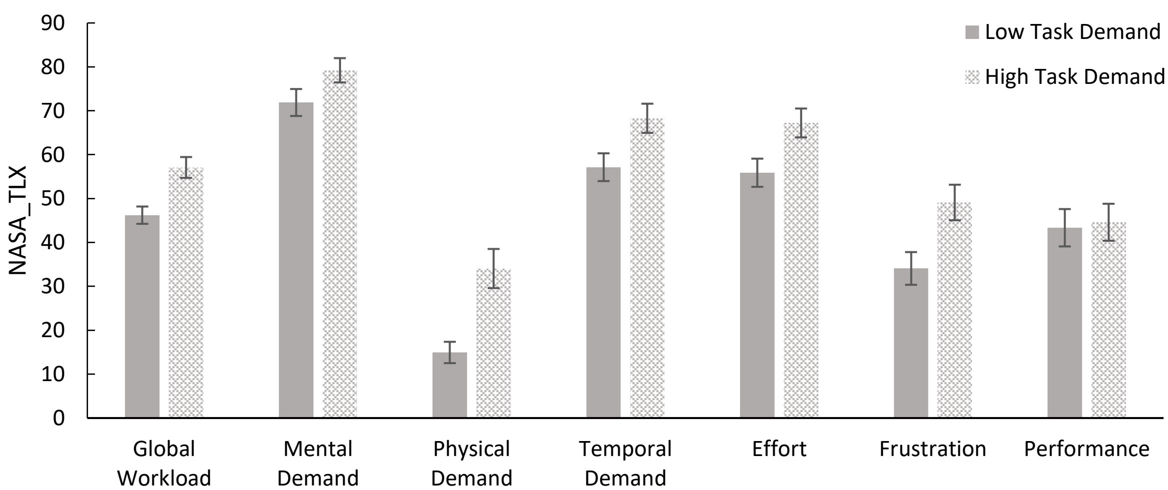

Figure 4. NASA-TLX workload factor ratings in low/high task demand conditions. Error bars represent standard errors.

Bonferroni-corrected *t*-tests were also computed to check the impact of LOA manipulations. Mean differences are shown in Figure 5. No significant self-rated workload differences were found between different LOA conditions. Therefore, the following analyses will focus on the impact of task demand manipulations.

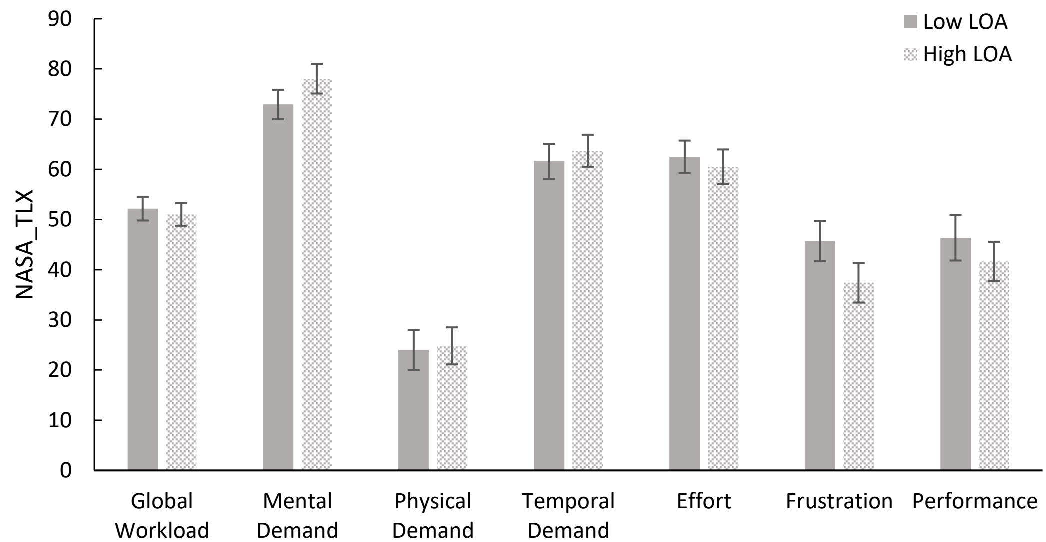

Figure 5. NASA-TLX workload factor ratings in low/high LOA conditions. Error bars represent standard errors.

Stress State

A series of 2 × 2 × 2 (LOA × task demand × pre- vs. post-task) mixed-model ANOVAs were run for each stress state factors, including task engagement, distress, and worry, to test the effects of experimental manipulations on subjective states. The results from ANOVA for task engagement showed a near significant interaction between pre-/post-task and task demand, $F(1, 97) = 3.65$, $p = .059$, $\eta^2_p = .04$ (Figure 6). In the low task demand condition, participants were less engaged after tasks, compared to the pre-task baseline. There was another significant interaction between pre-/post-task and task demand for distress, $F(1, 97) = 7.81$, $p < .01$, $\eta^2_p = .07$ (Figure 7). In the high task demand condition, participants reported greater distress after task exposure, compared to the pre-task baseline. Regarding worry, a significant main effect for pre-/post-task was found, $F(1, 97) = 46.14$, $p < .01$, $\eta^2_p = .32$ (Figure 8). Worry decreased in all conditions, and worry was lower in low task demand than in high task demand conditions.

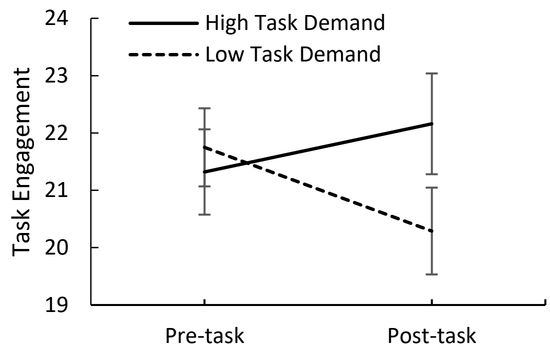

Figure 6. Pre- to post-task change in task engagement for different task demand conditions.
Error bars represent standard errors.

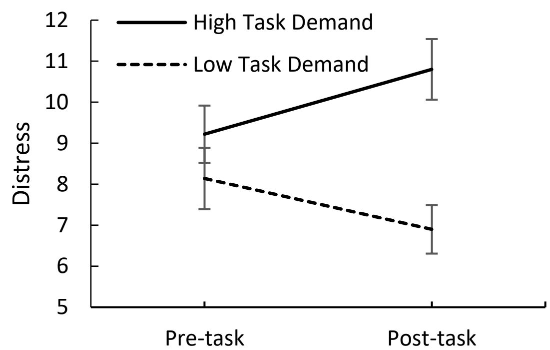

Figure 7. Pre- to post-task change in distress for different task demand conditions.
Error bars represent standard errors.

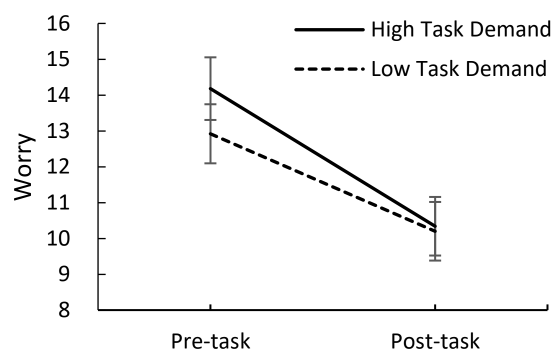

Figure 8. Pre- to post-task change in worry for different task demand conditions.
Error bars represent standard errors.

The Impact of LOAs and Task Demand on Task Performance

Three performance metrics for the two high priority surveillance tasks, Image Analysis and Weapon Release authorization, were analyzed. *Accuracy* was defined as the percentage of correct responses. *Reliance* was defined as the percentage of trials on which the participant followed the recommendation from the automation. *Neglect* was defined as the frequency of items that appeared in the task window but were not opened by the participant. Detailed performance metric formulas for Image Analysis and Weapon Release authorization tasks are listed in Table 5 and Table 6. The possible types of response are categorized as shown in Table 4.

Table 4
The possible types of response in two surveillance tasks

	Correct Answer	Incorrect Answer
Agree with Automation	Hit	Near Miss Far Miss
Disagree with Automation	Correct Rejection	False Alarm

Near Miss: within one of the correct answer
Far Miss: greater than one of the correct answer, only in Image Analysis task
True Miss: task timed-out, only in low LOA condition

Table 5
Performance metrics in the Image Analysis task

	Formula
Low LOA	
Accuracy	$\frac{Hit + CorrectRejection}{Hit + CorrectRejection + NearMiss + FarMiss + FalseAlarm + TrueMiss} \times 100\%$
Reliance	$\frac{Hit + NearMiss + FarMiss}{Hit + CorrectRejection + NearMiss + FarMiss + FalseAlarm + TrueMiss} \times 100\%$
Neglect	Number of tasks which the participant never opened
High LOA	
Accuracy	$\frac{Hit + CorrectRejection}{Hit + CorrectRejection + NearMiss + FarMiss + FalseAlarm} \times 100\%$
Reliance	$\frac{Hit + NearMiss}{Hit + CorrectRejection + NearMiss + FarMiss + FalseAlarm} \times 100\%$
Neglect	Number of tasks which the participant never opened

Table 6
Performance metrics in the Weapon Release authorization task

	Formula
Low LOA	
Accuracy	$\frac{Hit + CorrectRejection}{Hit + CorrectRejection + NearMiss + FalseAlarm + TrueMiss} \times 100\%$
Reliance	$\frac{Hit + NearMiss}{Hit + CorrectRejection + NearMiss + FalseAlarm + TrueMiss} \times 100\%$
Neglect	Number of tasks which the participant never opened
High LOA	
Accuracy	$\frac{Hit + CorrectRejection}{Hit + CorrectRejection + NearMiss + FalseAlarm} \times 100\%$
Reliance	$\frac{Hit + NearMiss}{Hit + CorrectRejection + NearMiss + FalseAlarm} \times 100\%$
Neglect	Number of tasks which the participant never opened

A series of 2 × 2 × 2 (LOA × task demand × task type) mixed-model ANOVAs were computed to test the impact of automation and workload on UAV operation performance.

Accuracy

For accuracy, participants performed less accurately in Weapon Release authorization task (M = 75.7) than Image Analysis task (M = 82.3), $F(1, 91) = 23.91$, $p < .01$, $\eta^2_p = .21$ (Figure 9). Another main effect of task demand was also significant for accuracy, $F(1, 91) = 5.87$, $p < .05$, $\eta^2_p = .06$. Participants in low task demand groups (M = 80.9) achieved greater accuracy than those in high task demand groups (M = 77.1) in the surveillance tasks. Accuracy in Weapon Release authorization task seemed to be more vulnerable to high task demand than Image Analysis task, even though the interaction between task type and task demand was not significant.

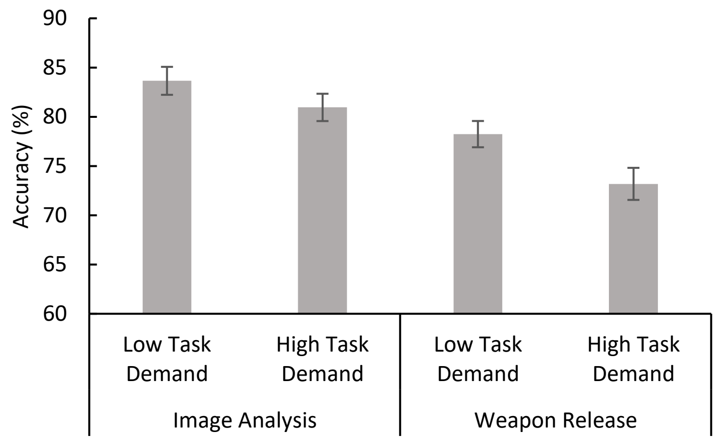

Figure 9. Task performance (accuracy) in the Image Analysis and the Weapon Release authorization tasks for different task demand conditions. Error bars represent standard errors.

Reliance on Automation

Reliance on automation was greater in the Image Analysis task (*M* = 75.6) than in the Weapon Release authorization task (*M* = 72.9), $F(1, 91) = 5.91$, $p < .05$, $\eta^2_p = .06$. A near significant main effect of task demand for reliance on automation was found, $F(1, 91) = 3.92$, $p = .051$, $\eta^2_p = .04$. Participants showed greater reliance on automation in low task demand conditions (*M* = 75.54) than in high task demand conditions (*M* = 72.98). Result also revealed a significant main effect of LOA for reliance, $F(1, 91) = 5.11$, $p < .05$, $\eta^2_p = .05$ (Figure 10). High LOA groups (*M* = 75.64) were more reliant on automation than low LOA groups (*M* = 72.76). In addition, the interaction between task type and task demand was also significant, $F(1, 91) = 4.76$, $p < .05$, $\eta^2_p = .05$ (Figure 11). In Weapon Release authorization task, task demand had a stronger effect on reliance on automation. Specifically, in Weapon Release authorization task, participants were less reliant on automation in high task demand conditions (*M* = 70.47) than in low task demand conditions (*M* = 75.38).

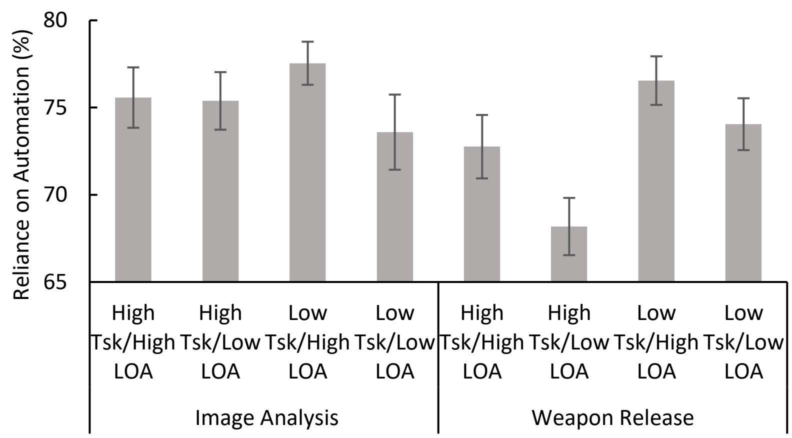

Figure 10. Task performance (reliance on automation) in the Image Analysis and the Weapon Release authorization tasks for different task demand and LOA conditions. Error bars represent standard errors.

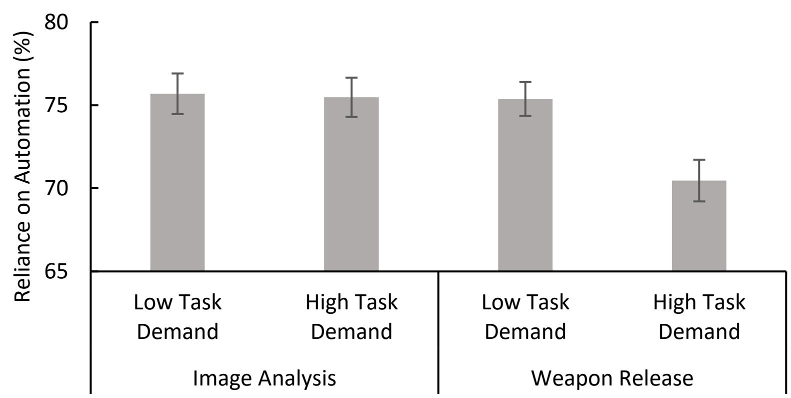

Figure 11. Task performance (reliance on automation) in the Image Analysis and the Weapon Release authorization tasks for different task demand conditions. Error bars represent standard errors.

Neglect

Regarding neglect, there was significantly more item neglects in Weapon Release authorization task (M = 8.9) than in Image Analysis task (M = 3.4), $F(1, 91) = 94.08$, $p < .01$, $\eta^2_p = .51$. The main effects for task demand and LOA were also significant for neglect (Figure 12). First, neglect was higher in high task demand groups (M = 8.4) than in low task demand groups (M = 3.9), $F(1, 91) = 19.18$, $p < .01$, $\eta^2_p = .17$. Second, neglect was higher in high LOA conditions (M = 7.1) than in low LOA conditions (M = 5.1), $F(1, 91) = 4.20$, $p < .05$, $\eta^2_p = .04$. In addition, the interaction between task type and task demand was significant, $F(1, 91) = 9.68$, $p < .01$, $\eta^2_p = .10$

(Figure 13). The effect of task demand had a stronger impact on the Weapon Release authorization task. Participants in the high task demand conditions neglected the most number of items (M = 12.01) in the Weapon Release authorization task.

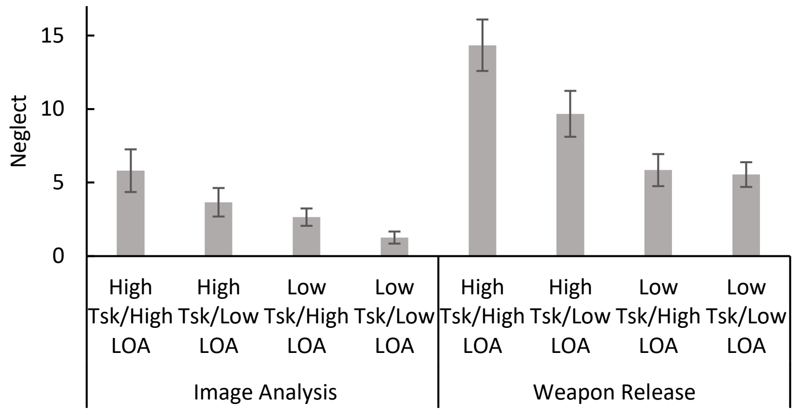

Figure 12. Task performance (neglect) in the Image Analysis and the Weapon Release authorization tasks for different task demand and LOA conditions. Error bars represent standard errors.

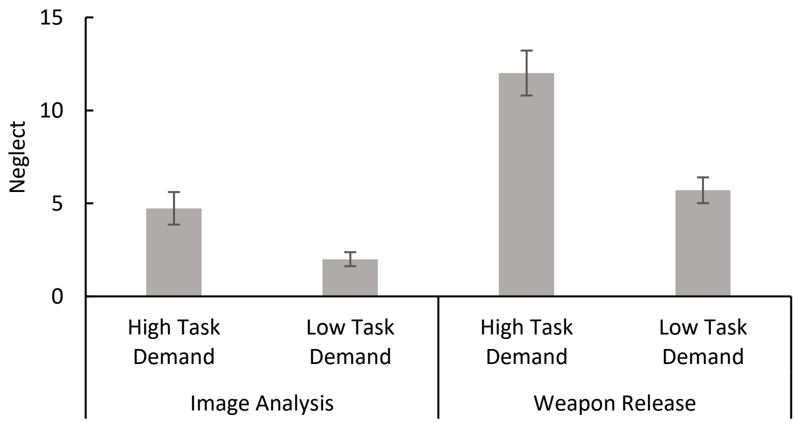

Figure 13. Task performance (neglect) in the Image Analysis and the Weapon Release authorization tasks for different task demand conditions. Error bars represent standard errors.

Individual Differences

Computer/Gaming Experience and Task Performance

Table 7 illustrates correlations between gaming experience and performance metrics in two surveillance tasks. Only one significant correlation was found for the Image Analysis task. Self-rated general computer expertise positively correlated with task accuracy (r = .246, p < .05).

For the Weapon Release authorization task, expertise in video games, first person shooter games, and other action games tended to be associated with higher accuracy, greater reliance on automation, and less neglect. Video game exposure time also showed the same trend of association with performance as the expertise factors. But the results were only significant for general video game exposure time and Weapon Release authorization task performance as well as between other action game exposure time and neglect in the Weapon Release authorization task.

Generally speaking, gaming experience, especially gaming expertise, was only predictive for the Weapon Release authorization task. The Weapon Release authorization task was rated more demanding than the Image Analysis task. In order to test whether task demand had a moderator effect, standardized gaming experience and task demand variables, as well as the interaction terms were added to the hierarchical regression models. The hierarchical regression results revealed that none of the tested interaction terms were significant. Therefore, task demand did not moderate the association between gaming experience and task performance.

Table 7
Correlations between gaming experience and performance metrics in the Image/Weapon Release tasks

	Image Analysis			Weapon Release		
	Accuracy	Reliance	Neglect	Accuracy	Reliance	Neglect
Computer daily hrs	-.074	-.025	.032	-.076	-.062	-.018
Computer expertise	.246*	.200	-.082	.166	.008	-.183
Game weekly hrs	.030	-.001	-.067	.235*	.249*	-.218*
Game expertise	.042	.070	.030	.293**	.270**	-.181
FPS weekly hrs	-.047	-.106	.049	.163	.129	-.112
FPS expertise	.043	-.016	-.032	.316**	.257*	-.252*
Action weekly hrs	-.024	-.068	-.090	.177	.191	-.228*
Action expertise	.061	.013	-.070	.369**	.331**	-.285**

** $p < .01$, * $p < .05$

Computer/Gaming Experience and Stress State

Computer and gaming experience correlated fairly consistently with more positive pre-task states (Table 8). Among those computer and gaming experience factors, computer expertise and action game expertise significantly correlated with all three pre-task state factors. Computer expertise significantly correlated with distress (r = -.334), task engagement (r = .329), worry (r = -.202). Action game expertise significantly correlated with distress (r = -.294), task engagement (r = .335), worry (r = -.216). Time spent on using computers or playing only showed positive relationships to pre-task task engagement, but no significant correlations with the other two state factors. Regarding the post-task state factors, only task engagement was positively related to time spent on using computers, playing video games, and first person shooter game expertise.

Table 8
Correlations between gaming experience and pre-/post-task stress state factors

	Pre-task			Post-task		
	Distress	Engagement	Worry	Distress	Engagement	Worry
Computer daily hrs	-.118	.282**	.012	-.111	.237*	.076
Computer expertise	-.334**	.329**	-.202*	-.167	.149	-.108
Game weekly hrs	-.164	.342**	-.148	-.167	.217*	-.125
Game expertise	-.293**	.296**	-.129	-.139	.078	-.100
FPS weekly hrs	-.158	.287**	-.158	-.079	.230*	-.019
FPS expertise	-.201*	.269**	-.160	-.120	.226*	-.132
Action weekly hrs	-.123	.199*	-.089	-.103	.078	-.086
Action expertise	-.294**	.335**	-.216*	-.167	.119	-.191

**p < .01, *p < .05

Gender Differences

Bonferroni-corrected t-tests were conducted to test gender differences in subjective stress state factors (Table 9) and objective performance metrics (Table 10). Women were significantly less engaged than men both before and after the tasks. Initially, women (M = 20.03, SD = 5.12) were less engaged than men (M = 23.64, SD = 4.15), t(97) = 3.91, p < .01. Levene's test indicated unequal variances, therefore the degrees of freedom were adjusted to 97.25. During the last 10 minutes of the experimental tasks, women (M = 20.10, SD = 6.14) were less engaged than men (M = 22.79, SD = 5.13), t(99) = 2.31, p < .05. In addition, women (M = 9.75, SD = 5.14) reported greater distress than men (M = 7.55, SD = 4.87) after the tasks, $t(99) = -2.17$, $p < .05$. Regarding the performance metrics, only one gender difference was found in terms of accuracy on the Weapon Release authorization task. Men (M = 79.52, SD = 7.64) performed more accurately than women (M = 73.13, SD = 11.44), $t(93) = 3.02$, $p < .01$.

Table 9
t-tests for gender differences in pre-/post-task stress state factors

	Male			Female			95% CI for Mean		
	M	*SD*	*n*	*M*	*SD*	*n*	Difference	*t*	*df*
Pre-task									
Distress	7.79	5.74	42	9.31	4.62	59	[-3.57, 0.53]	-1.47	99
Engagement	23.64	4.15	42	20.03	5.12	59	[1.78, 5.44]	3.91**	97.25
Worry	12.69	6.18	42	14.15	5.91	59	[-3.88, 0.95]	-1.20	99
Post-task									
Distress	7.55	4.87	42	9.75	5.14	59	[-4.21, -0.19]	-2.17*	99
Engagement	22.79	5.13	42	20.10	6.14	59	[0.38, 4.99]	2.31*	99
Worry	9.98	5.90	42	10.47	5.74	59	[-2.82, 1.83]	-.43	99

**$p < .01$, *$p < .05$

Table 10
t-tests for gender differences in performance metrics in the Image/Weapon Release tasks

	Male			Female			95% CI for Mean		
	M	*SD*	*n*	*M*	*SD*	*n*	Difference	*t*	*df*
Image									
Accuracy	83.54	8.32	38	81.45	10.53	57	[-1.96, 6.12]	1.02	93
Reliance	76.36	6.87	38	75.06	9.09	57	[-2.14, 4.74]	.75	93
Neglect	3.23	4.18	38	3.48	5.29	57	[-2.29, 1.77]	-.25	93
WR									
Accuracy	79.52	7.64	38	73.13	11.44	57	[2.18, 10.59]	3.02**	93
Reliance	74.48	7.31	38	71.84	8.72	57	[-0.76, 6.05]	1.54	93
Neglect	7.32	6.66	38	9.94	7.89	57	[-5.70, 0.47]	-1.68	93

**$p < .01$, *$p < .05$

In order to understand the gender differences better, gender differences in computer and gaming experience were tested using Bonferroni-corrected *t*-tests (Table 11). Although there was no gender difference in daily hours in using computers, women reported not only less time spent on playing video games, but less expertise in computer and video games. Specifically, comparing to men, women reported less general computer expertise, $t(66.95) = 3.48$, $p < .01$; less video game expertise, $t(98.74) = 8.10$, $p < .01$; less first person shooter game expertise, $t(99) = 9.25$, p

<.01; less other action game expertise, $t(99) = 7.84$, p <.01; less weekly game hours, $t(69.14) = 5.38$, p <.01; less weekly first person shooter game hours, $t(59.41) = 4.5$, p <.01; and less weekly other action game hours, $t(70.98) = 4.01$, p <.01. The degrees of freedom were adjusted because Levene's test results indicated the assumption of homogeneity of variances was violated.

Table 11
t-tests for gender differences in computer/gaming experience

	Male			Female			95% CI for Mean		
	M	*SD*	*n*	*M*	*SD*	*n*	Difference	*t*	*df*
Computer daily hrs	4.76	3.03	42	4.32	2.36	59	[-0.63, 1.50]	.81	99
Computer expertise	2.79	.75	42	2.32	.51	59	[0.20, 0.73]	3.48**	66.95
Game weekly hrs	3.90	2.18	42	1.85	1.39	59	[1.29, 2.82]	5.38**	64.19
Game expertise	5.07	1.20	42	2.81	1.60	59	[1.71, 2.81]	8.10**	98.74
FPS weekly hrs	2.81	1.89	42	1.36	1.06	59	[0.81, 2.10]	4.50**	59.41
FPS expertise	3.74	1.55	42	1.15	1.23	59	[2.03, 3.14]	9.25**	99
Action weekly hrs	2.98	1.88	42	1.61	1.38	59	[0.69, 2.05]	4.01**	70.98
Action expertise	4.69	1.65	42	2.20	1.52	59	[1.86, 3.12]	7.84**	99

**p < .01, *p < .05

In addition, in order to test if there was an association between gender and performance as well as subjective states with gaming experience controlled, multiple regressions were conducted, with Weapon Release accuracy as the dependent measure. The results indicated that, with relevant gaming experience factors, especially gaming expertise factors, controlled at the first step of the regression, gender predicted neither the subjective stress state factors nor objective performance accuracy. However, with gender entered at the first step, gaming experience factors remained predictive.

Task Performance and Stress State

Correlational analyses were computed to assess the relationship between task performance and pre-/post-task stress state factors. Pre-task worry was found to be the only factor that was negatively associated with task accuracy in both tasks (Image Analysis, $r = -.231$, $p < .05$; Weapon Release, $r = -.216$, $p < .05$) and reliance on automation in the Image Analysis task ($r = -.222$, $p < .05$). The correlational results were shown in Table 12.

Table 12
Correlations between performance metrics and pre-task stress state factors

	Pre-Distress			Pre-Engagement			Pre-Worry		
	Overall	Low	High	Overall	Low	High	Overall	Low	High
Image Analysis									
Accuracy	-.072	.034	-.140	.001	-.122	.111	-.231*	-.184	-.254
Reliance	-.065	.012	-.142	.073	-.008	.149	-.222*	-.212	-.232
Neglect	.129	-.012	.147	-.065	.182	-.164	.061	-.105	.091
Weapon Release									
Accuracy	-.171	-.158	-.129	.162	.279	.076	-.216*	-.165	-.224
Reliance	-.145	-.047	-.156	.048	.083	.017	-.165	-.123	-.153
Neglect	.191	.078	.189	-.189	-.179	-.220	.113	-.111	.180

***p* < .01, **p* < .05; Low: low task demand; High: high task demand

Significant correlations were found between all three post-task stress state factors and specific performance metrics (Table 13). Distress was associated with accuracy and neglect in both tasks, especially in high task demand conditions. Task engagement was negatively correlated with neglect in both tasks (Image Analysis, r = -.411, $p < .05$; Weapon Release, r = -.314, $p < .05$) in high task demand conditions. Worry showed a negative association with task accuracy, but this trend was only significant in the Image Analysis task when task demand was high (r = -.286, $p < .05$).

Table 13
Correlations between performance metrics and post-task stress state factors

	Post-Distress			Post-Engagement			Post-Worry		
	Overall	Low	High	Overall	Low	High	Overall	Low	High
Image Analysis									
Accuracy	-.268**	-.106	-.334*	.056	.028	.142	-.157	-.027	-.286*
Reliance	-.128	-.103	-.158	-.040	-.120	.037	-.142	-.050	-.235
Neglect	.303**	-.121	.334*	-.193	.045	-.411*	.138	-.088	.250
Weapon Release									
Accuracy	-.392**	-.211	-.408**	.049	-.032	.203	-.143	-.110	-.169
Reliance	-.207*	-.134	-.086	-.103	-.096	-.006	-.096	.037	-.201
Neglect	.382**	-.082	.408**	-.155	-.216	-.314*	.113	.018	.175

***p* < .01, **p* < .05; Low: low task demand; High: high task demand

Personality and Stress State

Table 14 shows the correlations between personality and stress state factors. Pre-task distress was correlated with all five personality factors. Among those, conscientiousness and neuroticism were significantly associated with pre-task distress in both low and high task demand conditions. Conscientiousness and neuroticism also were associated with pre-task engagement in high task demand group and in data pooled across task demand conditions (shown in "overall" columns).

Compared with the correlations in the pre-task states, the correlations between personality and post-task states showed the similar trend, but were generally weaker (Table 15). Among the five personality factors, conscientiousness seemed to be the most predictive one. Conscientiousness was significantly negatively associated with post-task distress in the low task demand condition and across conditions, and positively associated with post-task engagement in the high task demand condition.

Table 14
Correlations between personality factors and pre-task stress state factors

	Pre-Distress			Pre-Engagement			Pre-Worry		
	Overall	Low	High	Overall	Low	High	Overall	Low	High
Extraversion	-.263**	-.374**	-.159	-.026	.059	-.103	.092	.139	.027
Agreeableness	-.250*	-.357*	-.143	.120	.246	.019	.024	.007	.066
Conscientiousness	-.373**	-.432**	-.304*	.259**	.198	.308*	.007	-.057	.080
Neuroticism	.491**	.364**	.599**	-.229*	-.049	-.356*	.283**	.268	.279
Openness	-.266**	-.235	-.347*	.202*	.191	.237	.076	.205	-.117

***p* < .01, **p* < .05; Low: low task demand; High: high task demand

Table 15
Correlations between personality factors and post-task stress state factors

	Post-Distress			Post-Engagement			Post-Worry		
	Overall	Low	High	Overall	Low	High	Overall	Low	High
Extraversion	-.129	-.234	-.133	-.045	.071	-.184	.100	.023	.183
Agreeableness	-.242*	-.310*	-.130	-.005	-.003	.036	-.009	-.083	.053
Conscientiousness	-.269**	-.331*	-.205	.190	-.065	.430**	-.023	-.066	.020
Neuroticism	.213*	.126	.206	.156	.216	.082	.176	.317*	.066
Openness	-.110	-.096	-.223	.074	.020	.125	.016	.023	.004

***p* < .01, **p* < .05; Low: low task demand; High: high task demand

Task Demand as a Moderator between Personality and Stress State

Hierarchical regression analyses were conducted to test the potential moderator effect of task demand between personality and stress state. Relevant personality factors and task demand were standardized and regressed onto stress state in the first step, followed by the interaction term of personality and task demand in the second step. Variance inflation factor (VIF) values did not imply any multicollinearity issue in the tested regression models.

A single moderator effect of task demand was found between conscientiousness and post-task engagement, $\beta = .25$, $t(97) = 2.64$, $p = .01$. The positive association between conscientiousness and post-task engagement was stronger when task demand was higher. Simple slopes analysis indicated that although there was a positive association between conscientiousness and post-task engagement when task demand was high (1 standard deviation above mean), $\beta = 2.58$, $t(97) = 3.16$, $p < .01$, this association was not present when task demand was low (1 standard deviation below mean), $\beta = -.36$, $t(97) = -.48$, $p = .64$ (Figure 14).

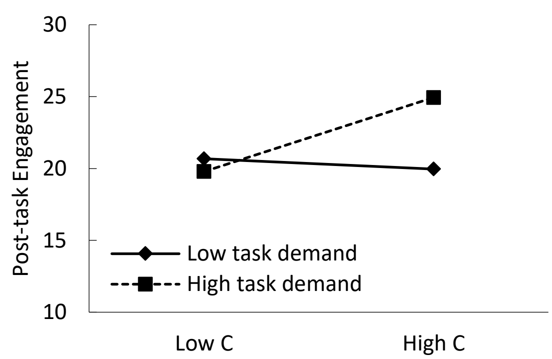

Figure 14. Association between conscientiousness (C) and post-task engagement moderated by task demand.

Personality and performance

Table 16 shows the correlations between personality and Image Analysis task performance in different task demand conditions. Personality did not predict performance on the Image

Analysis task very well. Only conscientiousness was found to be negatively associated with reliance on automation in the high task demand condition (r = -.353, p < .05), and across conditions (r = -.219, p < .05).

Table 16
Correlations between personality factors and performance metrics in the Image Analysis task

	Accuracy			Reliance			Neglect		
	Overall	Low	High	Overall	Low	High	Overall	Low	High
Extraversion	.095	.156	.056	.095	.089	.102	-.087	-.201	-.088
Agreeableness	.102	-.083	.192	-.055	-.057	-.062	-.138	.169	-.166
Conscientiousness	-.176	-.226	-.155	-.219*	-.079	-.353*	-.186	-.084	-.220
Neuroticism	.004	-.092	.125	.032	-.118	.156	-.018	.117	-.138
Openness	.009	-.039	.082	-.159	-.260	-.024	-.028	-.124	.003

**p < .01, *p < .05; Low: low task demand; High: high task demand

Table 17 shows the correlation between personality and Weapon Release authorization task performance in different task demand conditions. Personality factors showed an opposite tendency in predicting reliance on automation and neglect in different task demand conditions. Conscientiousness was negatively associated with reliance on automation in high task demand condition (r = -.372, p < .01), but tended to be positively associated with reliance on automation in low task demand condition (r = .064, p = .67). Additionally, conscientiousness was negatively correlated with neglect in high task demand condition (r = -.285, p < .05), but tended to be positive correlated with neglect when task demand was low (r = .112, p = .46). Also, a negative correlation was found between agreeableness and neglect in the high task demand condition (r = -.298, p < .05), but the correlation tended to be positive, though nonsignificant, in the low task demand condition (r = .172, p = .25). Besides these, agreeableness was also associated with reliance on automation in high task demand condition (r = -.345, p < .05). Extraversion was negatively correlated with reliance on automation in the high task demand condition (r = -.339, p < .05) and across conditions (r = -.225, p < .05).

Table 17
Correlations between personality factors and performance metrics in the Weapon Release task

	Accuracy			Reliance			Neglect		
	Overall	Low	High	Overall	Low	High	Overall	Low	High
Extraversion	-.079	.027	-.136	-.225*	-.056	-.339*	-.044	.017	-.146
Agreeableness	-.093	.019	-.246	-.166	-.065	-.345*	-.229*	.172	-.298*
Conscientiousness	-.038	.147	-.210	-.156	.064	-.372**	-.165	.112	-.285*
Neuroticism	-.022	-.238	.183	-.058	-.159	.090	.021	-.083	-.052
Openness	.032	-.078	.170	-.009	-.193	.204	-.117	-.253	-.080

**$p < .01$, *$p < .05$; Low: low task demand; High: high task demand

Task Demand as a Moderator between Personality and Performance

Hierarchical regression analyses were conducted to test the potential moderator effect of task demand between personality and performance. Relevant personality factors and task demand were standardized and regressed onto performance metrics in the first step, followed by the interaction term of personality and task demand in the second step. Variance inflation factor (VIF) values did not imply any multicollinearity issue in the tested regression models.

Results confirmed that task demand moderated the association between conscientiousness and reliance on automation, $\beta = -.22$, $t(91) = -2.25$, $p < .05$. The negative association between conscientiousness and reliance on automation was stronger when task demand was higher. Simple slopes analysis indicated that although there was a negative association between conscientiousness and reliance on automation in high task demand condition (1 standard deviation above mean), $\beta = -3.11$, $t(91) = -2.67$, $p < .01$, this association was not present in low task demand condition (1 standard deviation below mean), $\beta = .46$, $t(91) = .43$, $p = .67$ (Figure 15).

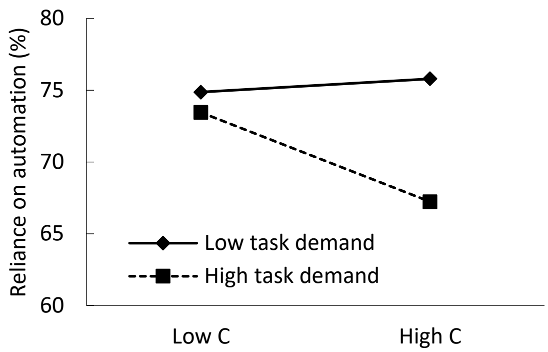

Figure 15. Association between conscientiousness (C) and reliance on automation in the Weapon Release authorization task moderated by task demand.

Results confirmed that task demand moderated the association between conscientiousness and neglect in the Weapon Release authorization task, $\beta = -.19$, $t(91) = -2.05$, $p < .05$. The negative association between conscientiousness and neglect was stronger in high task demand condition. Simple slopes analysis indicated that although there was a negative association between conscientiousness and neglect in the high task demand condition (1 standard deviation above mean), $\beta = -2.29$, $t(91) = -2.45$, $p < .05$, there was no such association in the low task demand condition (1 standard deviation below mean), $\beta = .55$, $t(91) = .54$, $p = .59$ (Figure 16).

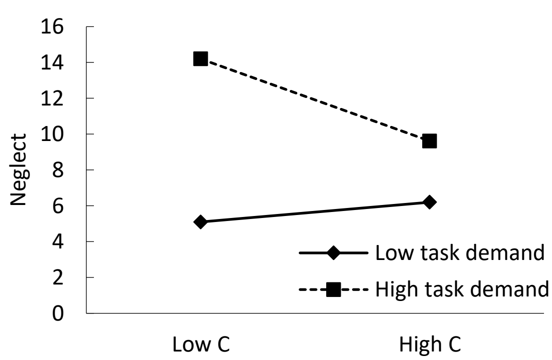

Figure 16. Association between conscientiousness (C) and neglect in the Weapon Release authorization task moderated by task demand.

Multiple regression analyses also indicated that task demand moderated the association between agreeableness and neglect in the Weapon Release authorization task, $\beta = -.21$, $t(91) = -2.18$, $p < .05$. The negative association between agreeableness and neglect was stronger in the high task demand condition. Simple slopes analysis indicated that there was a near significant negative association between agreeableness and neglect in high task demand condition (1 standard deviation above mean), β = -2.21, $t(91)$ = -1.97, p = .05; such an association was not observed in low task demand condition (1 standard deviation below mean), β = 1.00, $t(91)$ = 1.14, p = .27 (Figure 17).

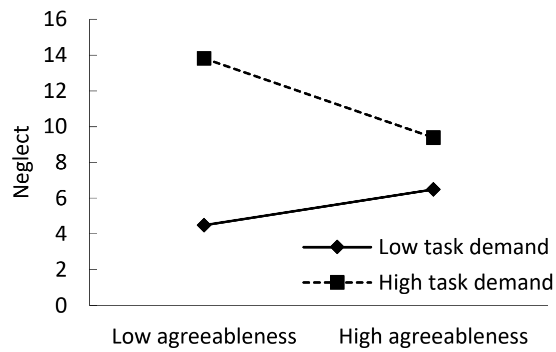

Figure 17. Association between agreeableness and neglect in the Weapon Release authorization task moderated by task demand.

Trust and Reliance on Automation

Subjective Trust on Surveillance Tasks

Participants' feedback after the experiments suggested that the reliability of the two surveillance tasks may be perceived as being at different levels, although the two tasks were set with same reliability (correct 80% of the time). A paired-samples t-test was run to compare subjective trust on the two tasks. Results indicated that there was no statistical difference between the subjective trust on Image Analysis task (M = 3.12, SD = .76) and Weapon Release authorization task (M = 3.08, SD = .77), $t(100)$ = .45, p = .66.

A further 2 × 2 × 2 (LOA × task demand ×task type) mixed-model ANOVA was computed to determine if there were group differences between LOA and task demand manipulations. Results confirmed that there was no difference in subjective trust between LOA conditions, $F(1, 97) = .95$, $p = .33$, $\eta^2_p = .01$; or between task demand conditions, $F(1, 97) = .08$, $p = .77$, $\eta^2_p = .00$.

Subjective Trust and Performance

Table 18 shows correlations between subjective trust and performance on the two surveillance tasks. Results suggested that there was no association between subjective trust and task performance.

Table 18
Correlations between subjective trust and performance metrics in the Image Analysis/Weapon Release tasks

	Image Analysis			Weapon Release		
	Accuracy	Reliance	Neglect	Accuracy	Reliance	Neglect
HC trust	-.135	-.052	.000	-.069	-.110	-.014
IM trust	.024	.123	-.118	.074	.050	-.173
WR trust	-.176	-.157	.064	.037	.089	-.033

HC: Human-computer; IM: Image Analysis task; WR: Weapon Release authorization task
$^{**}p < .01$, $^{*}p < .05$

Subjective trust and personality

Table 19 displays correlations between subjective trust and personality factors. Results suggested that there was no association between subjective trust and personality.

Table 19
Correlations between subjective trust and personality

	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness
HC trust	-.015	.043	.190	.071	-.106
IM trust	.010	.062	.121	.010	-.029
WR trust	.159	.063	.179	.080	.004

HC: Human-computer; IM: Image Analysis task; WR: Weapon Release authorization task
$^{**}p < .01$, $^{*}p < .05$

Subjective Trust and Gaming Experience

Table 20 illustrates correlations between subjective trust and gaming experience factors. Generally, the correlations were weak. Only two associations were statistically significant. Trust in the Image Analysis task and weekly time spent on playing other action games were positively correlated (r = .213, p < .05). Trust in the Weapon Release authorization task automation and general game expertise were positively correlated (r = .229, p < .05).

Table 20
Correlations between subjective trust and gaming experience

	Game Daily Hrs	Game Expertise	FPS Wkly Hrs	FPS Expertise	Action Wkly Hrs	Action Expertise
HC trust	-.028	.043	-.131	-.136	-.006	-.049
IM trust	.097	.153	-.040	-.075	.213*	.090
WR trust	.088	.229*	-.012	.048	.157	.154

HC: Human-computer; IM: Image Analysis task; WR: Weapon Release authorization task
**p < .01, *p < .05

DISCUSSION

A main objective of this study was to demonstrate that a multi-UAV simulation environment could be used to induce high workload and stress in participants. Supporting this objective, the task demand and LOA manipulations influenced task performance generally as expected. The task demand manipulations elicited higher subjective distress and workload. LOAs did not affect operator workload, but affected reliance behavior.

Another aim was to identify individual difference factors associated with performance and stress, in higher and lower task demand conditions. A variety of factors were associated with task performance and with subjective stress response. Video gaming experience was linked to lower distress and better performance, suggesting possible transfer of skills. Some gender differences were revealed in stress response and task performance, but all the gender effects became insignificant with gaming experience controlled. Generally, the effects of personality were consistent with previous studies, except for some novel findings with the performance metrics. Performance was negatively correlated with distress, consistent with previous research showing that distress impairs multi-tasking.

Personality may become more important for outcomes when the operator is challenged by high demands. The study confirmed that task demand seemed to moderate the influence of personality factors on stress response and performance metrics. Individuals high in conscientiousness and agreeableness tended to be more resistant to overload under high task demand circumstances. However, conscientiousness was associated with suboptimal use of automation under high demands.

Previous research has assumed that subjective trust mediates the impact of system reliability on reliance behavior. However, no significant correlation was found between subjective trust and reliance on automation. In addition, personality did not predict trust on automation in the UAV context either, suggesting that assessment of subjective trust is of limited utility in this context.

Overall, study findings have several implications for the human factors of UAV operations. Automation allowed even novices to perform quite well in a sensor operator role, but participants were challenged by the more difficult ISR task (Weapon Release). Under high task demands, detection performance was impaired, reliance on the automation declined, and participants were prone to neglect the task. Given that the automation was quite reliable, the decline in reliance is concerning, and shows disuse of automation when it is most needed. Analyses of individual differences suggested benefits to recruiting action video gamers, as well as individuals able to maintain states of task engagement and low distress during operations. Personality impacted reliance more strongly than performance accuracy. In particular, highly conscientious individuals were especially prone to show under-reliance under the most demanding conditions, suggesting a misplaced motivation to take control personally. Training solutions to performance vulnerabilities might focus on high-demand task configurations, taking into account the individual's dispositions.

Thus, this research effort provides a better understanding of the impact of automation and workload on human performance and stress in the UAV context. Study findings show that both objective performance and subjective stress responses are influenced by multiple tasks and personal factors in the multi-UAV environment. Providing appropriate support for the operator, including optimizing the use of automation, requires an understanding of how individual differences interact with task demands. The remainder of this discussion reviews theoretical and practical implications of the results, and suggests how limitations of this study might be addressed in future research.

The Impact of Task Demand on Subjective States and Performance

In this study, the level of task demand was successfully manipulated to simulate the task demand variation in UAV operations by configuring the frequency of secondary task events in the ALOA simulation. High task demand produced higher subjective workload and greater distress, confirming the task was stressful. This trend of elevated workload and distress is consistent with the finding in a previous UAV simulation study (Panganiban & Matthews, 2014). Worry was reduced relative to baseline in both task demand conditions. Typically, demanding tasks can induce decreases in worry, as attention is refocused from internal concerns to external demands (Matthews et al., 2013), as appears to be the case here. By contrast, low workload, monotonous UAV tasks may lead to mind-wandering, which may, in turn, contribute to the decreases in worry (Cummings et al., 2013). Also, the present result was consistent with the trend of greater declines in worry in high event rate vigilance tasks (Shaw et al., 2010).

Generally, in terms of accuracy and neglect, participants' performance was better in low task demand conditions. In high task demand conditions, less accuracy and more neglect were observed. The findings confirmed the hypothesis of the detrimental effects of high task demand on performance. Reliance on automation in the Image Analysis task was consistent across task demand conditions, while significantly less reliance on automation in the Weapon Release task was observed in the high task demand condition. Weapon Release was generally more difficult than Image Analysis. The lowest level of accuracy, reliance on automation, and the most instances of neglect in the Weapon Release task in the high task demand condition indicate the vulnerability of this task configuration to impairment in performance. High task demand was assumed to contribute to stress. The elevation of distress and workload suggested that the high task demand mission was indeed stressful as expected. Participants under such high task demand, especially when working on the demanding tasks, may become overloaded and fail to maintain their performance. In the view of CCM (Hockey, 1997), the high neglect under high task demand circumstances may suggest strategy changes, such as using avoidance coping and deliberately setting a lower task goal. The maintenance of task engagement across time may indicate that although the task was stressful, there was no loss of attentional resources associated with cognitive fatigue. The lower task demand condition was assumed to be potentially monotonous and fatiguing. It is hypothesized that participants under low task demand would show more reliance on automation. However, this main effect was only marginal. Additionally, no significant loss of task engagement was observed, suggesting fatigue was generally minimal. Although the simulated UAV tasks require attentional resources, their somewhat challenging and interesting task components may help to motivate operators to maintain engagement. Such features make the tasks differ from typical vigilance tasks, which are usually more monotonous.

The Impact of LOAs on Subjective States and Performance

Utilizing automation can reduce workload and enable single operators to manage multiple UAVs at the same time, but it may also introduce human factors issues, such as a loss of situation awareness and complacency issues (Endsley, 1996; Miller & Parasuraman, 2007). Two intermediate levels of automation were employed in the experimental manipulations. Some impact of the two LOAs was found.

The LOA manipulations did not affect subjective workload nor stress response. Higher LOA should have helped to reduce operator workload. A possible explanation is that two intermediate levels, management-by-consent and management-by-exception, were selected from the LOAs model (Parasuraman et al., 2000) in the study. These two levels were possibly too close to make a profound difference in the effect of LOA on workload and stress response. Alternatively, at the higher LOA, the operator may have reallocated attention to additional activities, such as secondary tasks, so that workload remained constant.

Even though no effect of LOA on subjective workload and stress states was found, the two LOA configurations succeeded in producing differences in task performance. Greater reliance on automation and more neglect were observed in higher LOA conditions (management-by-exception). Higher LOA may lead to a loss of situation awareness associated with vigilance decrement and complacency issues (Endsley, 1996; Endsley & Kiris, 1995; Miller & Parasuraman, 2007) and may, in turn, result in the observed greater reliance on automation and more neglect. No significant difference in task accuracy was found between LOA conditions. This may suggest that considering the automation is relatively reliable, LOAs only have a subtle effect on the overall accuracy even though higher LOAs encourage operators to rely on the automation more. Also, the two LOAs were at intermediate levels close to each other in Parasuraman's model (2000). Future study may test the trend in other LOAs. In summary, the hypothesis is partially confirmed, with higher operator reliance on automation when using higher LOAs, but not better performance, in terms of task accuracy.

Individual Differences

It is important to identify individual differences in performance and stress in order to determine which operators have the highest aptitude for multi-UAV control and to support operators that have specific vulnerabilities to suboptimal performance or stress. Different types of individual differences had an impact on operator performance in different ways. Multiple individual differences factors, including gaming experience, personality traits, gender, and subjective stress states were involved in the present study. It is likely that these different factors overlap and interact with one another. For example, there are gender differences in personality, such as higher neuroticism in women (Costa, Terracciano, & McCrae, 2001; McCrae & Terracciano, 2005), that may be associated with greater stress vulnerability. Personality might also influence interest in video gaming (Mehroof & Griffiths, 2010; Walther, Morgenstern, & Hanewinkel, 2012). A full investigation of such interdependencies was beyond the scope of this dissertation, but the dependence of gender differences in video gaming experiences was specifically investigated. Men are known to have greater exposure to gaming (Desai, Krishnan-Sarin, Cavallo, & Potenza, 2010), and increasing recruitment of women is a significant issue for the USAF. Otherwise, the key inter-relationships between individual difference factors, stress states, and performance are discussed separately, in the sections that follow.

Gaming and Performance

Gaming experience was predictive of both lower subjective stress state and higher performance in Weapon Release task. Gaming experience, especially self-rated expertise on general video games, first person shooter games, or other action games, was associated with superior task performance in the more demanding Weapon Release task. Participants reporting more expertise on video gaming showed greater accuracy, more reliance on automation, and less task neglect in the demanding task. It seemed that gaming expertise factors were more reliable predictors of performance than game exposure factors. The results were consistent with the advantages of experienced video gamers shown in previous simulated UAV studies (Cummings et al., 2010; McKinley et al., 2011). Besides gaming expertise, weekly hours spent on playing video games was also associated with performance in the Weapon Release task. Considering video game exposure is positively associated with sensory, perceptual, and attentional abilities (Spence & Feng, 2010), this finding may suggest that practice in video gaming may improve such abilities and skills, which may transfer to and benefit UAV operations. Notably, video games often require particular skills and techniques for allocating attention across multiple subtasks, which may also be necessary for UAV operations. A game such as *Call of Duty* requires the player to monitor other units placed in multiple locations in the screen display, requiring spatial attention, and to determine which possible action is of highest priority at any given time, requiring executive processing. Plausibly, such attentional skills generalize to ALOA, which also requires scanning multiple windows and prioritizing different subtasks.

However, individuals with aptitudes for acquiring attentional skills and may be more likely to be self-selected to play action video games. Perhaps, the positive associations of gaming experience and UAV task performance are due to the self-selection for attentional abilities. Further work is necessary to confirm that that gaming skills transfer directly to the multi-UAV context.

Gaming and Stress State

Gaming experience was positively associated with task engagement and negatively associated with both distress and worry before the task exposure. Experienced video gamers may be more confident about performing the complex UAV tasks. Therefore, participants with more gaming experience tended to feel less stress and more enjoyment prior to the task. Only a few gaming experience factors, such as time spent on using computers, playing video games, and playing first person shooter games, were positively correlated with post-task engagement. In terms of post-task distress and worry, the same trends as pre-task were reported, but were not statistically significant. The associations between gaming experience and positive subjective states indicated that experienced video gamers may experience higher self-efficacy, which keeps them engaged in the tasks. Contrary to the negative stereotype described by Chappelle et al. (2014), gamers were no more stress-prone than those lacking gaming experience, and actually sustained task engagement more effectively over time. Similar to the associations with task performance, gaming expertise factors were more predictive of operator stress state than was gaming experience.

Stress State and Performance

Performance was associated with both pre-task and post-task stress state measures from the DSSQ. Performance correlates of pre-task measures indicate that states can predict future performance, which may be important for application. However, measures taken post-task, that ask how the person felt during the task, may be more representative of the states actually experienced during performance.

Pre-task worry was predictive of poor performance in terms of accuracy on both surveillance tasks. Evidence from previous studies indicated that worry may slow switching tasks (Johnson, 2009) and predict poor vehicle control (Funke et al., 2007). Consistent with the detrimental effects on performance, worry impaired UAV operation as well, suggesting that worry may impair the temporary resource availability for information processing, as attention is diverted from the task to processing personal concerns. Pre-task worry also predicted less reliance on automation, with a stronger effect in the Image Analysis task. Worry may impair attention, working memory, and executive control of multi-tasking (Matthews & Campbell, 2010; Matthews et al., 2013; Zeidner, 2010), and in turn, impair operator performance in the UAV context. The association between worry and performance became weaker after task exposure. Accuracy in the Image Analysis task remained significantly correlated with worry under high task demand.

High post-task distress correlated with poorer performance on both tasks in terms of accuracy and neglect. These associations were much stronger under high task demand. Distress response is primarily driven by subjective workload. High task demand manipulation can produce large amounts of workload. Similar negative associations between distress and performance were also seen in previous vigilance studies (Matthews, Hancock, & Desmond, 2012; Shaw et al., 2010), and on a dual-tasking working memory task (Matthews & Campbell, 2010). Attention Control Theory (ACT) argues that anxiety may interfere with executive control, and specifically the inhibition of task-irrelevant stimuli (Eysenck & Derakshan, 2011). Distress was also found to be associated with poor inhibition of task-irrelevant stimuli (Matthews & Zeidner, 2012). Thus, while distress may produce some general impairment in focused attention, its further association with impaired executive control may be especially damaging to performance in multi-tasking environments such as ALOA, where strategic deployment of attention across the different task windows is critical.

Post-task engagement was found to be positively associated with task performance in terms of neglect, but higher engagement was unrelated to accuracy. Task engagement is typically associated with superior executive control and reflects effort committed to achieving task goals, as well as higher overall resource availability (Matthews & Zeidner, 2012; Matthews et al., 2002). Higher engagement is also associated with performance in some applied settings, such as superior vehicle control in a moderately fatiguing simulated driving context (Funke et al., 2007). In the present data, remarkably, the association between task engagement and performance differs from the typical association in vigilance tasks. The task engagement-performance correlation is typically around 0.3 (Matthews et al., 2013), whereas no significant correlation was found between task engagement and task accuracy in both ISR tasks. This suggested that UAV operations may require different information processing mechanisms to vigilance. Overall resource availability may not be critical for ISR accuracy, although resource shortfalls may become more important when operators are fatigued and lose task engagement. Task engagement effects may have reflected motivation rather than resource availability. Under high task demands, it is difficult to maintain attention to all the various subtasks. Consistent with Hockey's (1997) theory that fatigue lowers task goals, low-engagement participants may reduce effort and neglect more ISR missions, while high-engagement participants may be better able to maintain effort and have less neglect. The present study did not assess stress process such as appraisal and coping, but previous studies suggest that appraising the task as challenging is critical for maintaining task engagement (Matthews et al., 2013; Saxby et al., 2013), and the high engagement operators here may have appraised maintaining high performance on all task elements as a motivating challenge.

Personality and Performance

Generally, there was no association between personality and task accuracy. Previous research (e.g., Finomore, Matthews, Shaw, & Warm, 2009) has found that correlations between attentional tasks and major personality factors such as the Big Five tend to be rather task-specific and relatively small in magnitude. The configurations of ALOA used here may not be conducive to demonstrating personality effects, although personality might be more predictive of accuracy under other circumstances.

Some correlations between personality and reliance on automation and neglect were found on both surveillance tasks. Specifically, extraversion was predictive of less reliance on automation in the Weapon Release task, especially under high task demand. Agreeableness and conscientiousness were also related to less reliance on automation in the Weapon Release task under high task demand. Because Weapon Release is more demanding than Image Analysis, these findings suggest that personality becomes increasingly predictive of reliance as demands increase. In addition, conscientiousness was negatively associated with reliance on automation in the Image Analysis task, especially under high task demand. These correlations between personality traits and reliance on automation were contrary to Szalma and Taylor's findings (2011), which identified no significant correlations between these personality traits and agreement with automation. Again, personality-performance associations may be somewhat task-specific. Because these associations depended on task demands, further discussion is reserved for the section on task moderator effects below.

Personality and Stress State

Neuroticism was predictive of less positive subjective states in advance of task performance. All the other four traits (extraversion, agreeableness, conscientiousness, and openness) were associated with less distress before the tasks. This trend was consistent but generally weakened after task exposure. These findings were consistent with the general trend in previous studies (Matthews, Joyner, Gilliland, et al., 1999; Matthews, Warm, Shaw, et al., 2010; Matthews et al., 2006; Shaw et al., 2010), for personality-stress correlations to attenuate over time, suggesting that personality may influence anticipation of stress more strongly than the actual experience of the task. Neurotic individuals tend to experience negative affective states such as anxiety, anger, and sadness due to perceived uncertainty of the task or to a tendency to appraise tasks as more threatening (Matthews et al., 2009). Therefore, they are more vulnerable to stress. In addition, individuals high in neuroticism may have more negative anticipation prior to the mission. On the contrary, other traits may promote a more pleasant mood, higher confidence, and lower tension, due to various biases in appraisal and coping (Matthews et al., 2013).

Gender

Similar to the previous findings of negative state (higher distress) and poorer performance in women in a simulated driving study (Matthews, Joyner, & Newman, 1999), some gender differences were found in stress response, and task performance in UAV operation. Initially, women were less engaged than men, but this effect attenuated toward the end of the task. Also, women reported greater distress after task exposure. In terms of task performance, women were less accurate in the more demanding Weapon Release task. No gender difference in reliance on automation was found on both tasks.

No gender difference was noticed in daily hours of using computers. However, women reported significantly less general computer expertise and gaming experience, including expertise and time spent on playing different kinds of video games, consistent with the trend of more gaming experience in men reported in previous surveys (Terlecki et al., 2011). All the gender differences in stress response and task performance became nonsignificant after gaming experience was controlled in the multiple regression models. This finding suggested that gender differences in stress response and performance may be side effects of the greater interest in gaming exhibited by men. Although men may have some high aptitudes in traditional military piloting, such as spatial processing (Carretta, 1997; Halpern, 2013), this may not generalize to UAV operations. The demands on spatial attention of the ISR tasks may differ from those of conventional flying. ALOA has a spatial component in that attention must be focused and refocused across multiple screen windows. However, there is little spatial uncertainty involved, and hence little need for visual search across the display for critical signals. The two primary surveillance tasks are demanding because of the similarity of the target and nontarget stimuli, not because of any difficulty in localizing stimuli in space. Video gaming may contribute to acquiring relevant skills, but gender does not seem to be, once gaming experience is controlled.

Task Demand as a Moderator

Associations between individual factors and performance during unmanned vehicle operations may vary in different task configurations (Szalma & Taylor, 2011). In the operational context, the operator's ability to deal with increases in task demand and overload may be critical for mission success, so moderator analyses here focused on the task demand manipulation. Task demand was found to moderate the impact of personality on stress response and task performance, as anticipated.

Individuals high in conscientiousness were more engaged under high task demand. Such an advantage was not observed when task demand was low. Also, high conscientiousness individuals tended to rely less on automation and show less neglect of ISR tasks under high task demand, whereas conscientiousness did not influence reliance behavior or neglect under low task demand. By successfully performing moderately challenging tasks, conscientious individuals may demonstrate self-efficacy and thrive in the tasks (Szalma & Taylor, 2011). High conscientiousness individuals may tend to take charge of controlling the task personally, instead of relying on automation, especially when the task is demanding and stressful. Such a strategy of taking control personally may lead to the observed high task engagement, low reliance on automation, and less neglect under high task demand. The elevation of task engagement experienced by more conscientious operators might confer both greater resource availability and stronger task motivation. However, motivation may be more important than resources for the observed impacts of conscientiousness, given the lack of association between task engagement and accuracy of task performance.

A moderator effect of task demands was also found for agreeableness and neglect in the Weapon Release task. High agreeableness individuals tended to have less neglect of tasks under high task demand, but more neglect of tasks under low task demand. As an interpersonal trait, agreeableness includes the propensity to trust others. To the extent that trust generalizes to automated systems, agreeable individuals may be less likely to misuse or disuse the automation. Also, agreeableness was found to correlate with less avoidant coping (Matthews & Campbell, 1998), which may discourage neglect in demanding conditions. High agreeableness individuals appeared to be more resistant to overload in challenging tasks. Again, motivational effects may be the predominant factor for less neglect in high task demand condition. In this case, motivations may be social in nature, such as complying with the experimenter's instructions, rather than linked to individual achievement as may be the case for conscientiousness.

Trust in Automation

Analyses of subjective trust confirmed no perceived difference in reliability between the two surveillance tasks, corresponding to the lack of objective difference. It is argued that trust is a mediator between reliability of automation and reliance on automation (Lee & See, 2004; Lee & Moray, 1992; Parasuraman & Wickens, 2008). Higher reliability of automation should induce greater trust, and in turn, elicit greater reliance on automation. Hence, it was expected that subject trust should correlate with reliance on automation. However, no significant correlation was found between subjective trust and reliance on automation or other performance metrics. In addition, there were no correlations between subjective trust and personality traits. Although agreeableness as an interpersonal trait is characterized by a propensity to trust others, it did not predict trust in automation, which was consistent with a previous simulated UGV study (Szalma & Taylor, 2011). By contrast, trust in the Weapon Release task automation was positively correlated with general video game expertise, and trust in the Image Analysis task was positively correlated with weekly hours in playing other action video games. But these associations were not consistent among other gaming experience factors. The mostly nonsignificant findings on the possible correlates of subjective trust suggested that subject trust does not necessarily directly affect the reliance behavior or operator performance. Personality traits may not play a critical role in sensitivity to the trustworthiness of the automation, but exposure to video games or other possible systems with automated components may have an impact on trust on automation. Video gamers may have some acquired insight into trust in computer systems. Generally, though, subjective trust does not seem to guide behavioral reliance. One possibility is that with an unfamiliar system, participants do not attend to their own subjective trust in making reliance decisions. Also, given that the time-pressured nature of the task gives little opportunity for reflecting on the behavior or the automation, “trust” in this context may be an unconscious process. Subjective trust may be a more meaningful metric in contexts where operators are familiar with the automation, but current findings suggest that researchers should be cautious about using subjective trust measures in laboratory studies of automated systems.

Limitations and Future Work

Firstly, the LOAs selected in this study were two intermediate levels in Parasuraman's 10-level LOA model (2000). These two selected LOAs did not make a profound difference in the impact of workload and stress response, although there was an effect on reliance, as anticipated. Higher LOA should be instrumental in reducing operator workload (Miller & Parasuraman, 2007), but no workload reduction was observed here. Future research may employ a wider range of LOAs and also test performance without automation support to investigate the impact of LOAs. However, human factors research may be most important with configurations such as the present one where both the automation and the human are fallible and optimization of reliance is critical.

Additionally, the reliability of the automation may influence operator's reliance behavior. Also, a previous study indicated that automation reliability may moderate the effect of personality traits, such as conscientiousness and openness, on operator performance and stress response in a simulated UGV task (Szalma & Taylor, 2011). Future research may utilize automation with different reliabilities to test its impact in UAV context, at intermediate LOAs.

Secondly, it was thought that the low task demand manipulation might induce passive fatigue in the form of large-magnitude declines in task engagement, as seen in automated vehicle driving studies (Saxby et al., 2013), and plausibly also during real-world monotonous UAV missions (Cummings et al., 2013). The one-hour duration of the task is sufficient to cause strong fatigue symptoms in vigilance studies (e.g., Shaw et al., 2010), but the loss of engagement in the low task demand condition here was minor. Possibly, the game-like task components in the UAV simulation helped to keep the participants engaged in the task and maintain their attention. Task duration was also considerably shorter than the missions often undertaken by operators. Future research may need to extend the mission durations and lower the task load to induce passive fatigue on operators.

Thirdly, participants in this study were of course much less trained than actual UAV operators. Adequate training may enable the operators to become more resistant to stress, although it might also reduce the sense of challenge which may have helped to sustain task engagement in the naïve student participants here. Future work should consider using more extensive comprehensive practice to ensure high levels of competence.

Finally, college students were recruited as participants in this study. The sample of college students may not represent the population of real military UAV operators, although USAF seeks to recruit from this potential pool of applicants. Thus, findings from this study may need to be confirmed by utilizing a sample of military personnel. Due to the limitation of the participant pool, more women were recruited than men in the study. The gaming experience was also not balanced with respect to gender. Future studies may balance the gender and gaming experience to disentangle the individual differences in these factors. Future research may also consider including psychophysiology measures, such as EEG, ERP, CBFV and eye tracking metrics, as of fatigue and trust in UAV operations. Psychophysiological assessments may be particularly useful if linked to reliance on automation, given that subjective trust measures were not predictive of reliance and individual differences may reflect unconscious processes.

PRACTICAL IMPLICATIONS

The demand for automated UAV support has been growing at an unprecedented rate in the military (Schanz, 2010). Although it can help to reduce physical threats to the aircraft (Gertler, 2012), augment surveillance and combat capabilities (Chappelle et al., 2010), and bring plenty of other benefits, there still remains some human factors issues. First, single operator control of multiple UAVs is anticipated to be a particularly time-critical, cognitively demanding multi-task work environment (Calhoun et al., 2011; Guznov et al., 2011). In response, developments are underway to extensively automate UAV functions with the goal of enhancing the operator's ability to manage task demands. However, rather than attempting to automate everything, and leave functions that cannot be reliably automated to the human, automation should be designed to support continual human engagement and maintained situation awareness (Eggers & Draper, 2006). Second, the current training pipeline for UAV operators cannot meet the growing demand (Paullin, Ingerick, Trippe, & Wasko, 2011). The growing demand may require extending the current recruitment population and improving current training effectiveness. Third, UAV operations involve considerable task demand variation which may be stressful and fatiguing. In control of multiple UAVs, the cost of task interruption and task switching may be particularly critical (Eggers & Draper, 2006). Therefore, it is important to monitor fatigue for testing fitness for duty prior to the task and checking capacity for continuing duty during a mission.

Design of Automated Systems

It is a significant priority for the USAF to effectively apply automation to future systems (Dahm, 2010). There remains a critical need for human involvement to facilitate successful UAV missions, especially in ISR missions which are often time critical and involve complex target, friendly, and non-combatant identification and discrimination (Eggers & Draper, 2006). Future missions may require one operator to control multiple UAVs. This study implies that an intermediate level of automation may be adequate for supporting operator performance of the task without excessive stress or fatigue. Performance deteriorated in the higher demand condition, but not catastrophically so, but additional operator supporter under high workload may be needed.

An Intermediate Level of Automation Can Aid Operator Performance

Performance on the ISR tasks was fairly good at both LOAs, with accuracy levels ranging from 75.7% - 80.9%. Higher accuracy would be required in an operational setting, but performance was adequate for a naïve sample given limited training. Performance data suggested that although LOA did not affect task accuracy directly, it had impacts on reliance on automation and neglect. Management-by-exception, the higher LOA, induced greater reliance on automation and more neglect in the surveillance tasks. Previously, variation in neglect was attributed to motivational factors, and the higher LOA may have had a demotivating effect on participants. Alternatively, the increased neglect may have resulted in a loss of situation awareness (Kaber & Endsley, 2004). The greater reliance on automation may indicate a misuse of automation, such as complacency issues. High reliability of the automation may contribute to the maintenance of task accuracy with increased neglect in the tasks. Neglect would be a concern in the operational environment because the automation cannot function until the operator initiates the mission. Management-by-consent may be the preferable intermediate LOA for aiding operator performance as well as helping to maintain situation awareness.

Demanding Tasks Need More Automation Aid

Generally, the more demanding task (Weapon Release authorization task) showed lower accuracy, less reliance on automation, and more neglect. Additionally, the Weapon Release task was particularly vulnerable to high task demand, in which condition it showed the lowest accuracy, least reliance on automation, and most neglect. Therefore, demanding tasks may call for more automation aid to optimize the performance. However, the tendency found here for increasing task demand to lower reliance on automation tends to negate the benefits of automation when it is most needed. As previously discussed, this effect may reflect the tendency of the operator to take charge personally when the task is perceived as maximally training. Also, the automation should be highly reliable. High reliability can enable the system to achieve the task goal, and if reliability is high enough operators may be more willing to trust the automation under the most demanding conditions.

Adaptive LOA May Mitigate Operator Fatigue

Diagnostic monitoring of operator state, discussed below, may support adaptive automation that allows the automation to compensate for performance vulnerabilities associated with excessive workload, stress, and fatigue (Kaber & Endsley, 2004). One form of compensation is to adjust the LOA upwards or downwards, depending on the specific vulnerability.

UAV operations usually feature considerable task demand variation. Decreased task engagement in the low task demand condition and increased distress in the high task demand condition were observed in this simulated UAV study. Decreased task engagement may be an indicator of the beginning of a passive fatigue state, although the effect was small in magnitude. Continuous monitoring of the operator’s state using psychophysiological sensors might be able to detect the onset of both overload/distress and loss of task engagement. Adaptive automation helps to enhance human-machine interaction and is necessary for effective performance and fault management in complex systems (Parasuraman, Mouloua, & Hilburn, 1999; Moray, Inagaki, &

Itoh, 2000). Typically, increased distress is primarily driven by excessive task demands. Automation that responds to signs of distress by elevating LOA may help to alleviate perceived workload and avoid excessive stress. However, the present results suggest that a switch from management-by-consent to management-by-exception may not be sufficient to mitigate distress and workload. A better approach might be to switch some task components to full automation so that the human operator can focus intensively on tasks beyond the capability of automated systems.

Conversely, passive fatigue might be countered by shifting to a lower LOA that enables the operator to gain more manual control of the system, and in turn, to reengage to the mission. Therefore, adaptive LOA may be beneficial to mitigate operator fatigue and optimize operator performance. As fatigue was minor in this study, it does not support detailed recommendations, but further research could explore whether management-by-consent is a low enough LOA to maintain task engagement, or whether the operator might need to take full control of some task components.

Personnel Selection and Training

The USAF has increasing needs for UAV operators. Currently, the majority of the UAV operators are recruited from officers with little flying experience who have completed a UAV training course (Paullin et al., 2011). This study may have some implications for extending the recruitment population and for designing more effective training methods.

Personnel Selection

No gender differences in task performance or stress response were found when gaming experience was controlled. In other words, although traditionally military pilots are mostly male, women did not show any basic disadvantage relative to men in this study. The USAF might thus make greater efforts to recruit female operators.

Even though gender did not play a vital role in task performance and stress response in UAV operation, men reported more experience and expertise in video gaming, and are more likely to self-identify as serious gamers (Terlecki et al., 2011). The role of gaming experience in the prediction of performance and stress response has some implications for selection of UAV operator. Experienced video gamers seemed to have better performance (greater accuracy and less neglect) and be less stress-prone in UAV operations (higher task engagement, lower distress, and worry). These advantages suggest that video gamers may have high level of specialized aptitudes, such as sensory, perceptual, and attentional abilities, for success in UAV operations. Directing recruitment towards gamers may thus be an effective strategy.

Given manpower shortages and increasing needs for UAV operators, recruitment of operators needs to be expanded from traditional groups to some new populations, such as women and video gamers. By contrast, personality data did not show any general performance deficits linked to the Big Five traits, which may limit their utility in selection. In terms of subjective outcomes, the association between neuroticism and post-task distress might suggest that, as in other potentially stressful work contexts (Matthews et al., 2009), highly neurotic individuals may not well be suited to UAV operation. Similarly, the high task engagement of conscientious individuals under high task demands suggests a possible benefit to recruiting these persons; high conscientiousness is beneficial to a variety of aspects of work behavior (Matthews et al., 2009).

Personnel Training

Understanding how the various individual difference factors relate to specific performance vulnerabilities may help to design personalized training directed towards the individual's weaknesses. For example, although video gamers showed better task performance in terms of accuracy and neglect, they seemed to place more trust in the automation in demanding tasks. Training may emphasize detrimental effects of misuse of automation to avoid over-reliance and complacency issues.

Based on the correlational analysis of personality and performance, extraversion, agreeableness, and conscientiousness predicted less reliance on automation in demanding tasks, especially under high task demand. Particularly, individuals high in conscientiousness seemed to be more engaged but were more reluctant to use automation aids in high task demand conditions. The negative association trend between conscientiousness and task accuracy in high task demand conditions suggested that managing the tasks manually under some circumstances may result in poor performance. An appropriate level of trust in the automation should be established and maintained in UAV operations (Lee & See, 2004).Therefore, training on how to calibrate trust to match the capabilities of the system is critical. Training operators to calibrate trust and use automation appropriately in demanding and stressful tasks is especially needed. Conscientious operators may need to learn to trust the automation under high demands, contrary to their inclination to take charge personally.

Additionally, individuals high in neuroticism seemed to have more negative anticipations prior to the mission and were more vulnerable to distress. Adequate training and practice prior to the real mission may help to eliminate negative expectancies and build confidence in personal proficiency.

Diagnostic Monitoring

As reported in surveys of UAV operators, fatigue overlaps with stress, which is complex and multifaceted (Ouma et al., 2011). Usually, fatigue and stress both may impair the capability of performing UAV missions. Diagnostic fatigue monitoring is vital for testing fitness for duty prior to the task and checking capacity for continuing of duty during a mission. Although some psychophysiological measures may be more applicable for real-time monitoring than subjective scales, this study using subjective measures provide a theoretical basis for future efforts at diagnostic monitoring.

Fitness for Duty

During protracted military operations, UAV operators may carry an increasing burden of stress and fatigue, especially when sleep must be curtailed. Fitness of duty testing may be performed to determine if the operator is ready to begin a work shift, or if they should rest. The subjective stress correlates of performance in this study suggest some strategies for assessment of fitness for duty.

Task engagement reflects effort committed to achieving task goals and a state of readiness for resource mobilization in task performance (Matthews et al., 2002; Matthews, Warm, Reinerman-Jones, et al., 2010). Distress is expected to be detrimental to the individual's attention, working memory, and multi-tasking (Matthews & Campbell, 2010). From the perspective of resource theory, task engagement represents the availability of a general attentional resource. Both task engagement and distress predict vigilance decrement (Shaw et al., 2010; Matthews et al., 2013). Findings of this study also support that the states of task engagement and distress are related to performance competence in the simulated UAV operations, especially when task demand is high. The stress response in training tasks such as the present simulated UAV operation may be diagnostic for the fitness for duty in real UAV missions.

Task engagement was correlated with superior performance on both surveillance tasks in terms of less neglect. The finding is consistent with the previous literature of task engagement predicting demanding task performance requiring attentional resources. For instance, high task engagement was predictive for superior control of the vehicle in a simulated driving study (Funke et al., 2007). Previously, the association between task engagement and lower neglect was attributed to motivational processes, but in more fatiguing task conditions, task engagement may be more generally predictive of attentional efficiency. Task engagement was found to be correlated with perceptual sensitivity and predict vigilance in multiple studies (Matthews, Davies, & Holley, 1990; Matthews et al., 1999; Langheim et al., 2007; Helton, Matthews, & Warm, 2009). Psychophysiological evidence also shows correlations between the state of task engagement, task-focused coping and right-hemisphere cerebral blood flow velocity (CBFV) measured by transcranial Doppler sonography (TCD) in predicting vigilance decrement in a vigilance task (Reinerman et al., 2006). In terms of coping processes, task engagement is most reliably associated with task-focused coping and less use of avoidance (Matthews et al., 2013). CBFV and EEG indices such as increased slow wave activity may be able to detect loss of task engagement (Matthews, Warm, Reinerman-Jones, et al., 2010), and so could be used to determine fitness for duty.

Distress was negatively correlated with performance on both surveillance tasks in terms of lower accuracy and more neglect. Usually, distress is primarily driven by workload on complex tasks (Matthews et al., 2013). Similar negative associations between distress and performance were also seen in previous vigilance studies (Matthews et al., 2012; Shaw et al., 2010). By contrast with task engagement predicting task performance requiring attentional resources, Matthews and Campbell (2010) found that distress was more predictive of the performance impairment on tasks requiring fewer demands on sustaining attention. Findings of associations between distress and poor inhibition of task-irrelevant stimuli (Matthews & Zeidner, 2012) support the suggestion that distress may interfere with executive control. In addition, distress is reliably associated with the use of emotion-focused coping in terms of coping strategy (Matthews et al., 2013). While stress is typically linked to autonomic arousal, further research is necessary to determine psychophysiological correlates of these psychological aspects of distress, which could then be used to determine if the operator was too distressed to perform effectively.

Ideally, subjective states of task engagement and distress could be employed as indices of fitness for duty. Task engagement may predict operator's attentional resource availability, an effort committed to achieving task goals, and use of positive coping strategies. Distress may reflect operator's vulnerability to workload in stressful tasks, and interference with executive control. However, operators may be motivated to conceal stress and fatigue in the real setting, limiting the ability of organizations to utilize the subjective states of task engagement and distress in training or simulated missions as an element of personnel fitness for duty checking procedures. Psychophysiological correlates of these states might serve instead to identify unfit operators, but further research is necessary to implement such a strategy.

Continuing Duty

UAV operations often feature long shift durations (Chappelle et al., 2011). Such prolonged UAV missions may deplete the pool of attentional resources due to operator stress and fatigue. Temporal performance decrement accompanied by increased subjective fatigue has been observed in previous studies (Harris, Hancock, & Harris, 2005; Lieberman et al., 2006). Monitoring changes in task engagement and distress – or rather their psychophysiological equivalents – may be diagnostic of harmful stress and fatigue states, and therefore, be helpful for diagnostic monitoring for fitness for continuing operator duties.

Prolonged UAV operations involve considerable workload variation, such as long periods of low workload and intense activities for brief periods (Cummings et al., 2007, 2013). Such workload variation may induce active or passive fatigue which are both detrimental to operator performance. Active fatigue is typically characterized by increased distress, whereas passive fatigue usually links to a loss of task engagement (Saxby et al., 2013; Matthews et al., 2013). Large-magnitude declines, typically greater than 1 standard deviation, in task engagement are often seen in passive fatigue manipulations (Saxby et al., 2008, 2013). Empirically, high workload can elevate distress easily. Increases in distress, sometimes exceeding 1 standard deviation, are commonly observed in high workload tasks (Matthews et al., 2013). However, instead of being driven directly by workload, the personal interpretations of workload and the coping strategies the person adopts may be more critical factors for driving distress.

According to the Compensatory Control Model (CCM; Hockey, 1997), active fatigue due to the stressor of high workload may produce “strain”, which may encourage operators to compensate for the impact of stress by increasing effort. Passive fatigue may be more detrimental to operator performance due to the loss of attentional resources (Warm et al., 2008) or strategic reduction in the allocation of effort (Hockey, 1997), such as less task-focused coping and lowering of performance goals. The onset of passive fatigue signaled by significantly increased distress may imply a possible deterioration of continuous duty before an actual performance decrement.

In sum, monitoring the state changes during missions may help to detect operator fatigue allowing for intervention prior to the actual performance decrement. Intervention might take the form of adaptive automation, as previously described, or actually pulling the operator from the work shift. However, measurement of subjective state changes during operations may be difficult and have other limitations, such as operators’ motivations to conceal stress. Psychophysiological indices, such as eye movement and cerebral blood flow velocity, may be tested in for their capacity to detect fatigue state and predict performance.

CONCLUSION

Operators were able to manage multiple UAVs and accomplish the simulated mission with the aid of automation at a fairly good though imperfect level of competence, even under high task demands. Although there were individual differences in stress response, reliance on automation, and task performance, this present work demonstrated the feasibility of a single operator managing multiple UAVs using different LOAs under different task demands. Future research and development on how to improve dynamic interfaces in UAV operation and optimize operator reliance on automation must be driven by a deeper understanding of how individuals interact with automated systems and task demand, as well as the nature of the workload operators experience during the task. In addition, the findings may provide implications for future personnel selection, such as recruitment of UAV operators from nontraditional populations including video gamers and women, and for training operators to optimize reliance and performance based on individual differences in personality. The findings also provide a means for diagnosis of readiness for duty and monitoring operator fatigue for interventions, although implementation may require a better understanding of physiological correlates of stress states.

APPENDIX A: DEMOGRAPHICS QUESTIONNAIRE

Demographics Questionnaire

Gender ______ Age ______ Major ___________________

1. Do you have normal/corrected vision?

   ☐YES ☐NO

2. Are you in your usual state of health physically?

   ☐YES ☐NO

3. If NO, please briefly explain:

   ☐____________________________________

4. How many hours of sleep did you get last night?

   ☐______ hours

5. Have you had any caffeine in the last 12 hours?

   ☐YES ☐NO

6. What is your occupation?

   ☐____________________________________

7. What is the highest level of education you have had?

   ☐Less than 4 yrs of college ☐Completed 4 yrs of college ☐Other

8. When did you use computers in your education? (*Circle all that apply*)

   ☐Grade School ☐Jr. High ☐High School

   ☐Technical School ☐College ☐Did Not Use

9. Where do you currently use a computer? (*Circle all that apply*)

   ☐Home ☐Work ☐Library ☐Other________ ☐Do Not Use

10. How many hours per day do you use a computer?

    ☐______ hours

11. Which of the following best describes your expertise with computers?

    ☐Novice ☐Average ☐Proficient ☐Expert

12. Estimate the **average number of hours per week** you have spent **playing all video games** within the **past two years (**e.g., PlayStation, Xbox, computer games)

0-1	2-4	5-7	8-10	11-13	14-16	17-19	20+
☐	☐	☐	☐	☐	☐	☐	☐

13. Estimate your **level of expertise** playing video games, in general

(*0 = no expertise, 1=novice, 3 = intermediate, 6 = expert*)

0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

14. Estimate **average** number **of hours per week** you have spent **playing 'First Person Shooter' video games** within the **past two years (**e.g., *Call of Duty*)

0-1	2-4	5-7	8-10	11-13	14-16	17-19	20+
☐	☐	☐	☐	☐	☐	☐	☐

15. Estimate your **level of** expertise in playing First Person Shooter games

(*0 = no expertise, 1=novice, 3 = intermediate, 6 = expert*)

0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

16. Which First Person Shooter game have you played **the most**? (You may enter 'None')

☐ ___________________________________

17. Estimate average **number of hours per week** you have spent **playing other action video games** within the **past two years** (i.e, not First Person Shooter - e.g., *Grand Theft Auto*)

0-1	2-4	5-7	8-10	11-13	14-16	17-19	20+
☐	☐	☐	☐	☐	☐	☐	☐

18. Estimate your level **of expertise** in playing other action video games

(*0 = no expertise, 1=novice, 3 = intermediate, 6 = expert*)

0	1	2	3	4	5	6
☐	☐	☐	☐	☐	☐	☐

19. Which action video game have you played **the most**? (You may enter 'None')

☐ ___________________________________

APPENDIX B: 40 MINI-MARKER PERSONALITY SCALE

40 Mini-Marker Personality Scale

Please use this list of common human traits to describe yourself as accurately as possible. Describe yourself as you see yourself at the present time, not as you wish to be in the future. Describe yourself as you are generally or typically, as compared with other persons you know of the same sex and of roughly your same age. Before each trait, please write a number indicating how accurately that trait describes you, using the following rating scale:

1	2	3	4	5	6	7	8	9
Extremely Inaccurate	Very Inaccurate	Moderately Inaccurate	Slightly Inaccurate	Neither Inaccurate Nor Accurate	Slightly Accurate	Moderately Accurate	Very Accurate	Extremely Accurate

____ **Bashful**	____ **Energetic**	____ **Moody**	____ **Systematic**
____ **Bold**	____ **Envious**	____ **Organized**	____ **Talkative**
____ **Careless**	____ **Extraverted**	____ **Philosophical**	____ **Tempermental**
____ **Cold**	____ **Fretful**	____ **Practical**	____ **Touchy**
____ **Complex**	____ **Harsh**	____ **Quiet**	____ **Uncreative**
____ **Cooperative**	____ **Imaginative**	____ **Relaxed**	____ **Unenvious**
____ **Creative**	____ **Inefficient**	____ **Rude**	____ **Unintellectual**
____ **Deep**	____ **Intellectual**	____ **Shy**	____ **Unsympathetic**
____ **Disorganized**	____ **Jealous**	____ **Sloppy**	____ **Warm**
____ **Efficient**	____ **Kind**	____ **Sympathetic**	____ **Withdrawn**

APPENDIX C: COMPLACENCY POTENTIAL RATING SCALE

Complacency Potential Rating Scale

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings **AT THE MOMENT**.

Extremely disagree = 0, Somewhat disagree = 1,

Neither disagree nor agree = 2, Somewhat agree = 3, Extremely agree = 4

1. I think automated medical devices like CT and MRI scans provide very reliable images for doctors to interpret.

   0 1 2 3 4

2. Automated devices used in medicine save time and money in the diagnosis and treatment of disease.

   0 1 2 3 4

3. If I need to have a tumor in my body removed, I would choose to undergo computer-aided surgery using laser technology because it is more reliable and safer than manual surgery.

   0 1 2 3 4

4. Automated devices used in modern aircraft, such as the automatic landing system, have made air journeys safer.

   0 1 2 3 4

5. ATMs provide a safeguard against the inappropriate use of an individual's bank account by dishonest people.

   0 1 2 3 4

6. Automated devices used in aviation and banking have made work easier for both employees and customers.

   0 1 2 3 4

7. Even though the automatic cruise control in my car is set at a speed below the speed limit, I worry when I pass a police radar speed trap in case the automatic control is not working properly.

   0 1 2 3 4

8. Manually sorting through card catalogues is more reliable than computer-aided searches for finding items in a library.

   0 1 2 3 4

9. I would rather purchase an item using a computer than have to deal with a sales representative on the phone because my order is more likely to be correct using the computer.

   0 1 2 3 4

10. Bank transactions have become safer with the introduction of computer technology for the transfer of funds.

    0 1 2 3 4

11. I feel safer depositing my money at an ATM than with a human teller.

    0 1 2 3 4

APPENDIX D: DSSQ — 3 STATE QUESTIONNAIRE

DSSQ — 3 State Questionnaire

Pre-Task Questionnaire

Instructions. This questionnaire is concerned with your feelings and thoughts at the moment. Please answer **every** question, even if you find it difficult. Answer, as honestly as you can, what is true of **you**. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you feel **AT THE MOMENT**. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually the best.

Date today.................... Time of day now....................

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings **AT THE MOMENT**.

Definitely false = 0, Somewhat false = 1,
Neither true nor false = 2, Somewhat true = 3, Definitely true = 4

1	I felt concerned about the impression I am making.	**0 1 2 3 4**
2	I felt relaxed.	**0 1 2 3 4**
3	The content of the task was dull.	**0 1 2 3 4**
4	I thought about how other people might judge my performance	**0 1 2 3 4**
5	I was determined to succeed on the task.	**0 1 2 3 4**
6	I felt tense.	**0 1 2 3 4**
7	I was worried about what other people think of me.	**0 1 2 3 4**
8	I thought about how I would felt if I were told how I performed	**0 1 2 3 4**
9	Generally, I felt in control of things.	**0 1 2 3 4**
10	I reflected about myself.	**0 1 2 3 4**
11	My attention was directed towards the task.	**0 1 2 3 4**
12	I thought deeply about myself.	**0 1 2 3 4**
13	I felt energetic.	**0 1 2 3 4**
14	I thought about things that happened to me in the past	**0 1 2 3 4**
15	I thought about how other people might perform on this task.	**0 1 2 3 4**
16	I thought about something that happened earlier today.	**0 1 2 3 4**
17	I found the task was too difficult for me.	**0 1 2 3 4**
18	I found it hard to keep my concentration on the task.	**0 1 2 3 4**
19	I thought about personal concerns and interests.	**0 1 2 3 4**
20	I felt confident about my performance.	**0 1 2 3 4**
21	I examined my motives.	**0 1 2 3 4**
22	I felt like I could handle any difficulties I encountered	**0 1 2 3 4**
23	I thought about how I have dealt with similar tasks in the past	**0 1 2 3 4**
24	I reflected on my reasons for doing the task	**0 1 2 3 4**
25	I was motivated to try hard at the task.	**0 1 2 3 4**
26	I thought about things important to me.	**0 1 2 3 4**
27	I felt uneasy.	**0 1 2 3 4**
28	I felt tired.	**0 1 2 3 4**
29	I felt that I could not deal with the situation effectively.	**0 1 2 3 4**
30	I felt bored.	**0 1 2 3 4**

POST-Task Questionnaire

Instructions. This questionnaire is concerned with your feelings and thoughts while you were performing the task. Please answer **every** question, even if you find it difficult. Answer, as honestly as you can, what is true of **you**. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you felt **WHILE PERFORMING THE TASK**. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually the best.

Date today.................... Time of day now....................

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings **WHILE PERFORMING THE TASK**.

Definitely false = 0, Somewhat false = 1,
Neither true nor false = 2, Somewhat true = 3, Definitely true = 4

1	I felt concerned about the impression I am making.	**0 1 2 3 4**
2	I felt relaxed.	**0 1 2 3 4**
3	The content of the task was dull.	**0 1 2 3 4**
4	I thought about how other people might judge my performance	**0 1 2 3 4**
5	I was determined to succeed on the task.	**0 1 2 3 4**
6	I felt tense.	**0 1 2 3 4**
7	I was worried about what other people think of me.	**0 1 2 3 4**
8	I thought about how I would felt if I were told how I performed	**0 1 2 3 4**
9	Generally, I felt in control of things.	**0 1 2 3 4**
10	I reflected about myself.	**0 1 2 3 4**
11	My attention was directed towards the task.	**0 1 2 3 4**
12	I thought deeply about myself.	**0 1 2 3 4**
13	I felt energetic.	**0 1 2 3 4**
14	I thought about things that happened to me in the past	**0 1 2 3 4**
15	I thought about how other people might perform on this task.	**0 1 2 3 4**
16	I thought about something that happened earlier today.	**0 1 2 3 4**
17	I found the task was too difficult for me.	**0 1 2 3 4**
18	I found it hard to keep my concentration on the task.	**0 1 2 3 4**
19	I thought about personal concerns and interests.	**0 1 2 3 4**
20	I felt confident about my performance.	**0 1 2 3 4**
21	I examined my motives.	**0 1 2 3 4**
22	I felt like I could handle any difficulties I encountered	**0 1 2 3 4**
23	I thought about how I have dealt with similar tasks in the past	**0 1 2 3 4**
24	I reflected on my reasons for doing the task	**0 1 2 3 4**
25	I was motivated to try hard at the task.	**0 1 2 3 4**
26	I thought about things important to me.	**0 1 2 3 4**
27	I felt uneasy.	**0 1 2 3 4**
28	I felt tired.	**0 1 2 3 4**
29	I felt that I could not deal with the situation effectively.	**0 1 2 3 4**
30	I felt bored.	**0 1 2 3 4**

APPENDIX E: METRICS FOR TRUST IN AUTOMATION

Metrics For Trust In Automation

1	**Completion of all tasks** was:	Very Difficult	Difficult	Moderately Easy	Easy	Very Easy
2	The **interfaces** to complete the tasks were:	Unacceptable	Bad	Satisfactory	Good	Optimum
3	To what extent was using the interfaces **frustrating**?	Not At All	A Little	Sometimes	Frequently	All the Time
4	My **performance** (all tasks) was:	Very Low	Low	Average	High	Very High
5	To what extent did you **trust the automation**?	No Trust	Low Trust	Some Trust	High Trust	Very High Trust
6	Rate your level of **workload**.	Bored	Somewhat Busy	Busy	Very Busy	Overloaded
7	To what extent was the **training & instructions** adequate?	Not At All	Somewhat	No Opinion	Pretty Much	Completely
8	To what extent is the **Router** competent in suggesting routes?	Not At All	A Little	Sometimes	Frequently	All the Time
9	To what extent can the **Router's** routes be predicted?	Not At All	A Little	Sometimes	Frequently	All the Time
10	To what extent can you rely on the **Router** to plan the routes?	Not At All	A Little	Sometimes	Frequently	All the Time
11	To what extent is the **Router** consistent in planning the routes?	Not At All	A Little	Sometimes	Frequently	All the Time
12	To what extent are you confident in the **Router's** performance?	Not At All	A Little	Sometimes	Frequently	All the Time
13	To what extent is the Automation competent **Counting Shapes**?	Not At All	A Little	Sometimes	Frequently	All the Time
14	To what extent is Automation predictable in **Counting Shapes**?	Not At All	A Little	Sometimes	Frequently	All the Time
15	To what extent can you rely on Automation in **Counting Shapes**?	Not At All	A Little	Sometimes	Frequently	All the Time
16	To what extent is the Automation consistent in **Counting Shapes**?	Not At All	A Little	Sometimes	Frequently	All the Time
17	To what extent are you confident in the Automation's performance **Counting Shapes**?	Not At All	A Little	Sometimes	Frequently	All the Time

18	To what extent is the Automation competent **<u>Detecting Targets</u>**?	Not At All	A Little	Sometimes	Frequently	All the Time
19	To what extent is Automation predictable in **<u>Detecting Targets</u>**?	Not At All	A Little	Sometimes	Frequently	All the Time
20	To what extent can you rely on Automation in **<u>Detecting Targets</u>**?	Not At All	A Little	Sometimes	Frequently	All the Time
21	To what extent is the Automation consistent in **<u>Detecting Targets</u>**?	Not At All	A Little	Sometimes	Frequently	All the Time
22	To what extent are you confident in the Automation's performance **<u>Detecting Targets</u>**?	Not At All	A Little	Sometimes	Frequently	All the Time

APPENDIX F: HUMAN-COMPUTER TRUST SCALE

Human-Computer Trust Scale

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings.

CONSIDER ONLY THE TRIAL YOU JUST COMPLETED!

Extremely disagree = 0, Somewhat disagree = 1,
Neither disagree nor agree = 2, Somewhat agree = 3, Extremely agree = 4

1. The automation responds the same way under the same conditions at different times.

   0 1 2 3 4

2. If I am not sure about a decision, I have faith that the automation will provide the best solution.

   0 1 2 3 4

3. The advice the automation produces is as good as that which a highly competent person could produce

   0 1 2 3 4

4. I understand how the automation will assist me with a decision I have to make.

   0 1 2 3 4

5. I can rely on the automation to function properly.

   0 1 2 3 4

6. I believe advice from the automation even when I don’t know for certain that it is correct.

   0 1 2 3 4

7. I like using the automation for decision making.

   0 1 2 3 4

8. Although I may not know exactly how the automation works, I know how to use it to make decisions.

   0 1 2 3 4

9. Overall, I trust the automation.

   0 1 2 3 4

APPENDIX G: NASA-TLX

INSTRUCTIONS: TLX RATINGS

We are interested in evaluating the experiences you had during the task. In the most general sense, we are examining the “workload” you experienced. The factors that influence workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by different task elements may change as you get more familiar with a task, perform easier or harder versions of it, or move from one task to another.

The following set of six rating scales was developed for you to use in evaluating your experiences during different tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales in the table, please ask the experimenter about it. It is extremely important that they be clear to you. You may keep the descriptions with you for reference during the experiment.

After performing the task, you will be presented with six rating scales. You are asked to evaluate the task by marking each scale at the point which matches your experience. Each line has two endpoint descriptors that describe the scale. You can place a cross on the line anywhere between the two endpoints. Note that “Performance” goes from “good” on the left to “bad” on the right. This order has been confusing for some people.

Please consider your responses carefully in distinguishing among different task conditions and consider each scale individually.

RATING SCALE DEFINITIONS		
Title	**Endpoints**	**Descriptions**
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, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	***Low/High***	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
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 these goals?
EFFORT	***Low/High***	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION	***Low/High***	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

MENTAL DEMAND

Low High

PHYSICAL DEMAND

Low High

TEMPORAL DEMAND

Low High

PERFORMANCE

Good Poor

EFFORT

Low High

FRUSTRATION

Low High

APPENDIX H: IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Gerald Matthews** and Co-PIs: **Lauren Reinerman, Rebecca Leis, Ryan Wohleber**

Date: **April 01, 2015**

Dear Researcher:

On 4/1/2015, the IRB approved the following human participant research until 03/31/2016 inclusive:

Type of Review:	IRB Continuing Review Application Form Expedited Review
Project Title:	Sustaining Performance in Simulation UAV Operation: Pilot Study
Investigator:	Gerald Matthews
IRB Number:	SBE-13-09562
Funding Agency:	AFOSR, University of Cincinnati
Grant Title:	
Research ID:	1055976

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30 days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu.

If continuing review approval is not granted before the expiration date of 03/31/2016, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Joanne Muratori

Signature applied by Joanne Muratori on 04/03/2015 05:08:47 PM EDT

IRB manager

REFERENCES

Abich, J., Matthews, G., & Reinerman-Jones, L. (2015). Individual differences in UGV operation: A comparison of subjective and psychophysiological predictors. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 741–745. doi: 10.1177/1541931215591174

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *302*(1110), 311–324. doi: 10.1098/rstb.1983.0057

Bailey, N. R., & Scerbo, M. W. (2007). Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*, *8*(4), 321–348. doi: 10.1080/14639220500535301

Calhoun, G. L., Ruff, H. A., Draper, M. H., & Wright, E. J. (2011). Automation-level transference effects in simulated multiple unmanned aerial vehicle control. *Journal of Cognitive Engineering and Decision Making*, *5*(1), 55–82. doi: 10.1177/1555343411399069

Carretta, T. R. (1997). Group differences on US Air Force pilot selection tests. *International Journal of Selection and Assessment*, *5*(2), 115–127. doi: 10.1111/1468-2389.00051

Chappelle, W., McDonald, K., & King, R. E. (2010). *Psychological attributes critical to the performance of MQ-1 Predator and MQ-9 Reaper U.S. Air Force sensor operators* (AFRL-SA-BR-TR-2010-0007). Air Force Research Lab Brooks City-Base TX Human Performance Wing (711TH).

Chappelle, W., Salinas, A., & McDonald, K. (2011). *Psychological health screening of remotely piloted aircraft (RPA) operators and supporting units* (AFRL-SA-WP-TR-2011-0002). School of Aerospace Medicine Wright Patterson AFB OH.

Chappelle, W., Swearingen, J., Goodman, T., Cowper, S., Prince, L., & Thompson, W. (2014). *Occupational health screenings of U.S. Air Force remotely piloted aircraft (drone) operators* (AFRL-SA-WP-TR-2014-0007). School of Aerospace Medicine, Wright-Patterson AFB, OH.

Costa, P., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322–331. doi: 10.1037/0022-3514.81.2.322

Cox, T., & Ferguson, E. (1991). Individual differences, stress and coping. In C. L. Cooper, & R. Payne (Eds.), *Personality and stress: Individual differences in the stress process* (pp. 7–30). Oxford, England: John Wiley & Sons.

Cummings, M. L., Brzezinski, A. S., & Lee, J. D. (2007). The impact of intelligent aiding for multiple unmanned aerial vehicle schedule management. Retrieved from http://dspace.mit.edu/handle/1721.1/90287

Cummings, M. L., Clare, A., & Hart, C. (2010). The role of human-automation consensus in multiple unmanned vehicle scheduling. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi: 10.1177/0018720810368674

Cummings, M. L., Mastracchio, C., Thornburg, K. M., & Mkrtchyan, A. (2013). Boredom and distraction in multiple unmanned vehicle supervisory control. *Interacting with Computers*, *25*(1), 34-47. doi: 10.1093/iwc/iws011

Dahm, W. (2010). *US Air Force chief scientist report on technology horizons: A vision for Air Force science & technology during 2010-2030* (AF/ST-TR-10-01). US Air Force.

De Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217. doi: 10.1016/j.trf.2014.06.016

Deptula, D., & Mathewson, E. (2009). *Air Force Unmanned Aerial System (UAS) Flight Plan 2009-2047*. Air Force Washington DC Director Intelligence Surveillance and Reconnaissance.

Desai, R., Krishnan-Sarin, S., Cavallo, D., & Potenza, M. (2010). Video-gaming among high school students: Health correlates, gender differences, and problematic gaming. *Pediatrics*, *126*(6), E1414–E1424.

Desmond, P. A., & Hancock, P. A. (2001). Active and passive fatigue states. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue* (pp. 455-465). Mahwah, NJ: Laurence Erlbaum Associates.

Desmond, P. A., & Matthews, G. (2009). Individual differences in stress and fatigue in two field studies of driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(4), 265–276. doi: 10.1016/j.trf.2008.12.006

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(4), 564–572. doi: 10.1518/001872007X215656

Drury, J. L., & Scott, S. D. (2008). Awareness in unmanned aerial vehicle operations. *The International C2 Journal*, *2*(1), 1–10.

Eggers, J. W., & Draper, M. H. (2006). Multi-UAV control for tactical reconnaissance and close air support missions: operator perspectives and design challenges. In *Proc. NATO RTO Human Factors and Medicine Symp. HFM-135*. NATO TRO, Neuilly-sur-Siene, CEDEX, Biarritz, France.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. doi: 10.1518/001872095779049543

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163–181). Hillsdale, NJ: Lawrence Erlbaum Associates.

Endsley, M. R., & Kiris, E. O. (1994). *Situation awareness in FAA airway facilities maintenance control centers (MCC): Final report*. Lubbock, TX: Texas Tech University.

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. doi: 10.1518/001872095779064555

Eysenck, M. W., & Derakshan, N. (2011). New perspectives in attentional control theory. *Personality and Individual Differences*, *50*(7), 955–960. doi: 10.1016/j.paid.2010.08.019

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York, NY: Plenum Press.

Finomore, V., Matthews, G., Shaw, T., & Warm, J. (2009). Predicting vigilance: A fresh look at an old problem. *Ergonomics*, *52*(7), 791–808. doi: 10.1080/00140130802641627

Flemisch, F. O., & Onken, R. (2000). *Detecting usability problems with eye tracking in airborne battle management support*. Universitaet der Bundeswehr Muenchen Neubiberg (Germany FR).

Funke, G., Matthews, G., Warm, J. S., & Emo, A. K. (2007). Vehicle automation: A remedy for driver stress? *Ergonomics*, *50*(8), 1302–1323. doi: 10.1080/00140130701318830

Gertler, J. (2012). *U.S. unmanned aerial systems*. Washington DC: Library of Congress, Congressional Research Service.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26–42. doi: 10.1037/1040-3590.4.1.26

Guznov, S., Matthews, G., Funke, G., & Dukes, A. (2011). Use of the RoboFlag synthetic task environment to investigate workload and stress responses in UAV operation. *Behavior Research Methods*, *43*(3), 771–780. doi: 10.3758/s13428-011-0085-9

Hake, D. F., & Schmid, T. L. (1981). Acquisition and maintenance of trusting behavior. *Journal of the Experimental Analysis of Behavior*, *35*(1), 109–124. doi: 10.1901/jeab.1981.35-109

Halpern, D. F. (2013). *Sex differences in cognitive abilities: 4th Edition*. New York, NY: Psychology Press.

Hancock, P. A., Desmond, P. A., & Matthews, G. (2012). Conceptualizing and defining fatigue. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The handbook of operator fatigue* (pp. 64–73). Surrey, England: Ashgate Publishing.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *31*(5), 519–537. doi: 10.1177/001872088903100503

Hanson, M. L., & Harper, K. A. (2000, July). *An intelligent agent for supervisory control of teams of uninhabited combat air vehicles (UCAVs)*. Paper presented at the Unmanned Systems 2000 Conference, Orlando, FL.

Harris, W. C., Hancock, P. A., & Harris, S. C. (2005). Information processing changes following extended stress. *Military Psychology*, *17*(2), 115–128. doi: 10.1207/s15327876mp1702_4

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 (Ed.), *Advances in Psychology* (Vol. 52, pp. 139–183). Amsterdam, the Netherlands: North-Holland.

Helton, W. S., Matthews, G., & Warm, J. S. (2009). Stress state mediation between environmental variables and performance: The case of noise and vigilance. *Acta Psychologica*, *130*(3), 204–213. doi: 10.1016/j.actpsy.2008.12.006

Helton, W. S., Matthews, G., & Warm, J. S. (2009). Stress state mediation between environmental variables and performance: The case of noise and vigilance. *Acta Psychologica*, *130*(3), 204–213. doi: 10.1016/j.actpsy.2008.12.006

Hockey, G. R. J. (1986). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance* (pp. 1–49). Oxford, England: John Wiley & Sons.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, *45*(1–3), 73–93. doi: 10.1016/S0301-0511(96)05223-4

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71. doi: 10.1207/S15327566IJCE0401_04

Johnson, D. R. (2009). Emotional attention set-shifting and its relationship to anxiety and emotion regulation. *Emotion, 9*(5), 681–690. doi: 10.1037/a0017095

Johnson, L. C. (1982). Sleep deprivation and performance. In W. B. Webb (Eds.), *Biological rhythms, sleep, and performance* (pp. 111–141). New York, NY: John Wiley.

Johnson, R., Leen, M., & Goldberg, D. (2007). *Testing Adaptive Levels of Automation (ALOA) for UAV Supervisory Control* (AFRL-HE-WP-TR-2007-0068). Air Force Research Laboratory. Wright-Patterson Air Force Base, OH

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. doi: 10.1002/prs.680160304

Kaber, D. B., & Endsley, M. R. (1999). Level of automation effects on telerobot performance and human operator situation awareness and subjective workload. In M. W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 165–170). Mahwah, NJ: Lawrence Erlbaum Associates.

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*(2), 113–153. doi: 10.1080/1463922021000054335

Langheim, L., Matthews, G., Warm, J. S., Reinerman, L. E., Shaw, T. H., Finomore, V. S., & Guznov, S. (2007, July). The long pursuit: In search of predictors of individual differences in vigilance. Paper presented at the Thirteenth Meeting of the International Society for the Study of Individual Differences, Giessen, Germany.

Langner, R., Steinborn, M. B., Chatterjee, A., Sturm, W., & Willmes, K. (2010). Mental fatigue and temporal preparation in simple reaction-time performance. *Acta Psychologica*, *133*(1), 64–72. doi: 10.1016/j.actpsy.2009.10.001

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, *46*(8), 819–834. doi: 10.1037/0003-066X.46.8.819

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York, NY: Springer Publishing Company.

Lee, J. D., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243–1270. doi: 10.1080/00140139208967392

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*(1), 153–184. doi: 10.1006/ijhc.1994.1007

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. doi: 10.1518/hfes.46.1.50_30392

Lieberman, H. R., Niro, P., Tharion, W. J., Nindl, B. C., Castellani, J. W., & Montain, S. J. (2006). Cognition during sustained operations: Comparison of a laboratory simulation to field studies. *Aviation, Space, and Environmental Medicine*, *77*(9), 929–935.

Liu, D., Wasson, R., & Vincenzi, D. A. (2009). Effects of system automation management strategies and multi-mission operator-to-vehicle ratio on operator performance in UAV systems. *Journal of Intelligent and Robotic Systems*, *54*(5), 795–810. doi: 10.1007/s10846-008-9288-4

Madsen, M., & Gregor, S. (2000). Measuring human-computer trust. *Proceedings of the 11th Australasian Conference on Information Systems*, 6–8.

Matthews, G. (2001). Levels of transaction: A cognitive science framework for operator stress. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 5–33). Mahwah, NJ, US: Lawrence Erlbaum Associates.

Matthews, G. (2016). Multidimensional profiling of task stress states for human factors: A brief review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(6), 801–813. doi: 10.1177/0018720816653688

Matthews, G., & Amelang, M. (1993). Extraversion, arousal theory and performance: A study of individual differences in the EEG. *Personality and Individual Differences*, *14*(2), 347–363. doi: 10.1016/0191-8869(93)90133-N

Matthews, G., & Campbell, S. E. (1998). Task-induced stress and individual differences in coping. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*(11), 821–825. doi: 10.1177/154193129804201111

Matthews, G., & Campbell, S. E. (2010). Dynamic relationships between stress states and working memory. *Cognition and Emotion*, *24*(2), 357–373. doi: 10.1080/02699930903378719

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., et al. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, *2*(4), 315–340. doi: 10.1037/1528-3542.2.4.315

Matthews, G., Davies, D. R., & Holley, P. J. (1990). Extraversion, arousal and visual sustained attention: The role of resource availability. *Personality and Individual Differences*, *11*(11), 1159-1173.

Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). *Human performance: cognition, stress, and individual differences*. Hove, England: Psychology Press.

Matthews, G., Deary, I. J., & Whiteman, M. C. (2009). *Personality Traits*. New York, NY: Cambridge University Press.

Matthews, G., & Desmond, P. A. (1998). Personality and multiple dimensions of task-induced fatigue: A study of simulated driving. *Personality and Individual Differences*, *25*(3), 443–458. doi: 10.1016/S0191-8869(98)00045-2

Matthews, G., Emo, A. K., Funke, G., Zeidner, M., Roberts, R. D., Costa Jr., P. T., & Schulze, R. (2006). Emotional intelligence, personality, and task-induced stress. *Journal of Experimental Psychology: Applied*, *12*(2), 96–107. doi: 10.1037/1076-898X.12.2.96

Matthews, G., Hancock, P. A., & Desmond, P. A. (2012). Models of individual differences in fatigue for performance research. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The handbook of operator fatigue* (pp. 155–170). Surrey, England: Ashgate Publishing.

Matthews, G., Joyner, L. A., Gilliland, K., Campbell, S. E., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state “Big Three.” In I. Mervielde, I. J. Dreary, F. DeFruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 335–250). Tilburg, the Netherlands: Tilburg University Press.

Matthews, G., Joyner, L. A., & Newman, R. (1999). Age and gender differences in stress responses during simulated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *43*(18), 1007–1011. doi: 10.1177/154193129904301802

Matthews, G., Szalma, J. L., Panganiban, A. R., Neubauer, C., & Warm, J. S. (2013). Profiling task stress with the Dundee Stress State Questionnaire. In L. Cavalcanti & S. Azevedo (Eds.), *Psychology of stress: New research* (pp. 49–90). Hauppauge, NY: Nova.

Matthews, G., Warm, J. S., Reinerman-Jones, L. E., Langheim, L. K., Washburn, D. A., & Tripp, L. (2010). Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *Journal of Experimental Psychology: Applied*, *16*(2), 187–203. doi: 10.1037/a0019572

Matthews, G., Warm, J. S., Shaw, T. H., & Finomore, V. S. (2010). A multivariate test battery for predicting vigilance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(14), 1072–1076. doi: 10.1177/154193121005401405

Matthews, G., & Zeidner, M. (2012). Individual differences in attentional networks: Trait and state correlates of the ANT. *Personality and Individual Differences*, *53*(5), 574–579. doi: 10.1016/j.paid.2012.04.034

McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, *88*(3), 547–561. doi: 10.1037/0022-3514.88.3.547

McKinley, R. A., McIntire, L. K., & Funke, M. A. (2011). Operator selection for unmanned aerial systems: Comparing video game players and pilots. *Aviation, Space, and Environmental Medicine*, *82*(6), 635–642. doi: 10.3357/ASEM.2958.2011

Mehroof, M., & Griffiths, M. D. (2010). Online gaming addiction: The role of sensation seeking, self-control, neuroticism, aggression, state anxiety, and trait anxiety. *Cyberpsychology, Behavior, and Social Networking*, *13*(3), 313–316. doi: 10.1089/cyber.2009.0229

Merlo, J. L., Wickens, C. D., & Yeh, M. (1999). *Effect of reliability on cue effectiveness and display signaling* (Tech. Report ARL-99-4/FED-LAB-99-3). Savoy, IL: University of Illinois: Aviation Research Lab.

Mikulka, P. J., Scerbo, M. W., & Freeman, F. G. (2002). Effects of a biocybernetic system on vigilance performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *44*(4), 654–664. doi: 10.1518/0018720024496944

Miller, C. A., & Parasuraman, R. (2003). Beyond levels of automation: an architecture for more flexible human-automation collaboration. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(1), 182–186. doi: 10.1177/154193120304700138

Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(1), 57–75. doi: 10.1518/001872007779598037

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied*, *6*(1), 44–58. doi: 10.1037/1076-898X.6.1.44

Mouloua, M., Gilson, R., & Hancock, P. (2003). Human-centered design of unmanned aerial vehicles. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *11*(1), 6–11. doi: 10.1177/106480460301100103

Mouloua, M., Gilson, R., Kring, J., & Hancock, P. (2001). Workload, situation awareness, and teaming issues for UAV/UCAV operations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(2), 162–165. doi: 10.1177/154193120104500235

Näätänen, R. (1973). The inverted-U relationship between activation and performance: A critical review. In S. Kornblum (Ed.), *Attention and Performance Vol 4* (pp. 4–155). New York, NY: Academic Press.

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(5), 734–746. doi: 10.1177/0018720811423261

Neubauer, C., Matthews, G., Saxby, D., & Langheim, L. (2011). Individual differences and automation choice in simulated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 1563–1567. doi: 10.1177/1071181311551326

Ouma, J. A., Chappelle, W. L., & Salinas, A. (2011). *Facets of occupational burnout among U.S. Air Force active duty and national guard/reserve MQ-1 Predator and MQ-9 Reaper operators* (AFRL-SA-WP-TR-2011-0003). School of Aerospace Medicine Wright Patterson AFB OH.

Panganiban, A. R., & Matthews, G. (2014). Executive functioning protects against stress in UAV simulation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 994–998. doi: 10.1177/1541931214581208

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. doi: 10.1177/0018720810376055

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced "complacency". *The International Journal of Aviation Psychology*, *3*(1), 1–23. doi: 10.1207/s15327108ijap0301_1

Parasuraman, R., Mouloua, M., & Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. In M. W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 119-123). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. doi: 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*, *30*(3), 286–297. doi: /10.1109/3468.844354

Parasuraman, R., & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 511–520. doi: 10.1518/001872008X312198

Paullin, C., Ingerick, M., Trippe, D. M., & Wasko, L. (2011). *Identifying best bet entry-level selection measures for US Air Force remotely piloted aircraft (RPA) pilot and sensor operator (SO) occupations* (No. HRRO-FR-11-64). Human Resources Research Organization, Alexandria VA.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1–2), 187–195. doi: 10.1016/0301-0511(95)05116-3

Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(4), 601–614. doi: 10.1518/hfes.45.4.601.27092

Riley, V. (1994). A theory of operator reliance on automation. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 8–14). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rouse, W. B., & Rouse, S. H. (1983). Analysis and classification of human error. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(4), 539–549. doi: 10.1109/TSMC.1983.6313142

Ruff, H. A., Narayanan, S., & Draper, M. H. (2002). Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles. *Presence: Teleoperators and Virtual Environments*, *11*(4), 335–351. doi: 10.1162/105474602760204264

Saxby, D. J., Matthews, G., Hitchcock, E. M., Warm, J. S., Funke, G. J., & Gantzer, T. (2008). Effect of active and passive fatigue on performance using a driving simulator. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52*(21), 1751–1755. doi: 10.1177/154193120805202113

Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of Experimental Psychology. Applied*, *19*(4), 287–300. doi: 10.1037/a0034386

Schanz, M. V. (2010). The indispensable weapon. *Air Force Magazine*, *93*(2), 32-36.

Scott, C. L. (1980). Interpersonal trust: A comparison of attitudinal and situational factors. *Human Relations*, *33*(11), 805–812. doi: 10.1177/001872678003301103

Selye, H. (1976). The stress concept. *Canadian Medical Association Journal*, *115*(8), 718.

Shaw, T. H., Matthews, G., Warm, J. S., Finomore, V. S., Silverman, L., & Costa Jr., P. T. (2010). Individual differences in vigilance: Personality, ability and states of stress. *Journal of Research in Personality*, *44*(3), 297–308. doi: 10.1016/j.jrp.2010.02.007

Sheridan, T. B., & Hennessy, R. T. (1984). *Research and modeling of supervisory control behavior: Report of a workshop.* Washington, DC: National Academy

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. MIT Man-Machine Laboratory, Cambridge, MA.

Singh, I. L., Molloy, R., & Parasuraman, R. (1993). Automation- induced “complacency”: development of the Complacency-Potential Rating Scale. *The International Journal of Aviation Psychology*, *3*(2), 111–122. doi: 10.1207/s15327108ijap0302_2

Spence, I., & Feng, J. (2010). Video games and spatial cognition. *Review of General Psychology*, *14*(2), 92–104. doi: 10.1037/a0019491

Stafford, S. C., Oron-Gilad, T., Szalma, J. L., & Hancock, P. A. (2004). Individual differences related to shooting performance, in a police night-training shooting exercise. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(9), 1131–1135. doi: 10.1177/154193120404800902

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, *48*(10), 1294–1313. doi: 10.1080/00140130500252990

Stulberg, A. N. (2007). Managing the unmanned revolution in the U.S. Air Force. *Orbis*, *51*(2), 251–265. doi: 10.1016/j.orbis.2007.01.005

Szalma, J. L. (2009). Individual differences in human–technology interaction: Incorporating variation in human characteristics into human factors and ergonomics research and design. *Theoretical Issues in Ergonomics Science*, *10*(5), 381–397. doi: 10.1080/14639220902893613

Szalma, J. L., Hancock, P. A., Dember, W. N., & Warm, J. S. (2006). Training for vigilance: The effect of knowledge of results format and dispositional optimism and pessimism on performance and stress. *British Journal of Psychology*, *97*(1), 115–135. doi: 10.1348/000712605X62768

Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: The five factor model of personality. *Journal of Experimental Psychology: Applied*, *17*(2), 71–96. doi: 10.1037/a0024170

Teo, G., & Szalma, J. L. (2011). The effects of task type and source complexity on vigilance performance, workload, and stress. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 1180–1184. doi: 10.1177/1071181311551246

Terlecki, M., Brown, J., Harner-Steciw, L., Irvin-Hannum, J., Marchetto-Ryan, N., Ruhl, L., & Wiggins, J. (2011). Sex differences and similarities in video game experience, preferences, and self-efficacy: Implications for the gaming industry. *Current Psychology*, *30*(1), 22–33. doi: 10.1007/s12144-010-9095-5

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. doi: 10.1016/0749-596X(89)90040-5

Tvaryanas, A. P., Thompson, W. T., & Constable, S. H. (2006). Human factors in remotely piloted aircraft operations: HFACS Analysis of 221 mishaps over 10 years. *Aviation, Space, and Environmental Medicine*, *77*(7), 724–732.

Walther, B., Morgenstern, M., & Hanewinkel, R. (2012). Co-occurrence of addictive behaviours: Personality factors related to substance use, gambling and computer gaming. *European Addiction Research*, *18*(4), 167–174. doi: 10.1159/000335662

Warm, J. S. (1993). Vigilance and target detection. In C. D. Wickens & B. M. Huey (Eds.), *Workload transition: Implications for individual and team performance* (pp. 139–170). Washington, DC: National Research Council.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 433–441. doi: 10.1518/001872008X312152

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & & D. R. Davies (Ed.), *Varieties of attention* (pp.63-102). San Diego, CA: Academic Press.

Wickens, C. D. (2000). *Imperfect and unreliable automation and its implications for attention allocation, information access and situation awareness* (ARL- 00-10/NASA-00-2). Savoy: University of Illinois, Aviation Research Lab.

Wickens, C. D., Dixon, S., Goh, J., & Hammer, B. (2005). *Pilot dependence on imperfect diagnostic automation in simulated UAV flights: An attentional visual scanning analysis*. Paper presented at the 13th International Symposium on Aviation Psychology, Oklahoma City

Wickens, C. D., Gempler, K., & Morphew, M. E. (2000). Workload and reliability of predictor displays in aircraft traffic avoidance. *Transportation Human Factors*, *2*(2), 99–126. doi: 10.1207/STHF0202_01

Wickens, C. D., & Hollands, J. G. (2000). Attention, time-sharing, and workload. In C. D. Wickens & J. G. Hollands (Eds.), *Engineering psychology and human performance* (pp. 439–479). Upper Saddle River, NJ: Prentice Hall

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. doi: 10.1002/cne.920180503

Zeidner, M. (2010). Test anxiety. In W. E. Craighead & C. B. Nemeroff (Eds.), *The corsini encyclopedia of psychology*. Danvers, MA: John Wiley & Sons.