AFRL-RH-WP-TR-2020-0021

AN INVESTIGATION OF APTITUDE REQUIREMENTS FOR HUMAN OPERATORS IN HUMAN-AUTOMATION INTERACTION

Michael F. Brady
Zakia S. Shermadou
Anthony Gibson
Infoscitex Corporation
4027 Colonel Glenn Highway, Suite 210
Dayton, OH 45431

Thomas R. Carretta
Air Force Research Laboratory
Collaborative Interfaces and Teaming Branch
Warfighter Interface Division
Wright-Patterson AFB, OH 45433

April 2020
Interim Report

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING
AIRMAN SYSTEMS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2020-0021 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signed//
THOMAS R. CARRETTA
Work Unit Manager
Collaborative Interfaces and Teaming Branch
Warfighter Interface Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

//signed//
TIMOTHY S. WEBB, Ph.D., DR-IV
Chief, Collaborative Interfaces and Teaming Branch
Warfighter Interface Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

//signed//
LOUISE A. CARTER, Ph.D., DR-IV
Chief, Warfighter Interface Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01/04/2020	Interim	01-October 2018 – 31-March-2020

4. TITLE AND SUBTITLE
An Investigation of Aptitude Requirements for Human Operators in Human-Automation Interaction

5a. CONTRACT NUMBER
FA8650-14-D-6500

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Michael Brady, Zakia S. Shermadou, Anthony Gibson, and Thomas R. Carretta

5d. PROJECT NUMBER

5e. TASK NUMBER
07

5f. WORK UNIT NUMBER
HOSA

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Infoscitex Corporation
4027 Colonel Glenn Highway
Suite 210
Dayton, OH 45431

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Warfighter Interface Division
Collaborative Interfaces and Teaming Branch
Wright-Patterson AFB, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHCC

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2020-0021

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A. Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
88ABW-2020-1501; Cleared 23 Apr 2020

14. ABSTRACT
The purpose of this effort was to identify individual differences that affect Airmen aptitude requirements for jobs involving autonomous systems and human-automation interaction (HAI). This was addressed in two stages. First, we conducted a literature review of task demands and operator states relevant to HAI. Based on this review, we propose a model for understanding performance as a composite of operator states, operator behaviors, and distal outcomes. Second, we conducted a meta-analysis of correlations between individual differences and the elements of this model. Results from the meta-analysis suggest cognitive skills such as working memory are important to performance in an HAI context. Results for personality suggest the presence of moderators. We conclude by identifying measurement challenges and research gaps.

15. SUBJECT TERMS
TAPAS, NCAPS, SDI, personnel evaluation, personality assessment.

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
Unclassified	Unclassified	Unclassified	SAR	51	Thomas R. Carretta 19b. TELEPHONE NUMBER *(Include Area Code)*

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

1.0 INTRODUCTION

The integration of autonomous unmanned systems into the joint force structure continues to change the nature of many military jobs (Fahey & Miller, 2017). As these systems become more sophisticated and autonomous, jobs involving high levels of human-automation interaction (HAI) might have aptitude requirements not measured by current entry-level test batteries. Thus, there is a need to identify aptitude requirements for such jobs as they exist today, and anticipate aptitude requirements for performance with future autonomous systems. This effort has the potential to reap the same benefits as past selection efforts: improved job performance, reduced training costs, improved person-job fit and retention of qualified personnel (Carretta & Ree, 2003).

2.0 BACKGROUND

2.1 Autonomous Systems in the United States Air Force (USAF)

Automation is often conceptualized as "technology that actively selects data, transforms information, makes decision, or controls processes" (Lee & See, 2004, p. 50). In recent years, the terms *autonomy* and *autonomous systems* have become more prevalent to describe systems capable of achieving mission goals independently and compensating for system failures without external intervention (USAF Office of the Chief Scientist, 2015).

Researchers and practitioners alike have embedded these systems into many job sectors including the medical field, the military, and the manufacturing sector. The nature and capabilities of automated agents have differed across past implementations, with various levels of automation and autonomy incorporated into jobs. Highly automated technologies are replacing low-automated technology and manual tasks (Defense Science Board, 2012). As this trend continues, the Department of Defense (DoD) envisions a future in which near fully-autonomous systems and human operators collaborate as teammates (Fahey & Miller, 2017).

2.1.1 HAI and Human-Machine Teaming (HMT)

HMT is an increasingly relevant concept with important implications for the success of autonomous systems. It is important to note that not all interactions between human operators and automation should be described as HMT. HMT implies a relationship that goes beyond controlling and supervising autonomous agents (Chen & Barnes, 2014). For instance, cases in which a human operator is merely supervising the automation, or assigning tasks to the autonomous system would not fall under HMT.

The literature often specifies that HMT involves elements of coordination, collaboration, and knowledge sharing between the human operator and the autonomous system. Lyons, Wynne, Mahoney, and Roebke (2019) reviewed the teaming literature to address the distinction between HMT and mere interaction. Characteristics of effective teaming include shared goals, shared awareness, desire for interdependence, motivation towards common objectives, action toward common objectives, and trust among team members (Groom & Nass, 2007; Lyons, Wynne, Mahoney, & Roebke, 2019).

Our distinction between HMT and HAI is not meant to imply that human operators on HMTs will not perform tasks associated with HAI (e.g., supervision and multitasking). Rather, such tasks might be performed in both in HAI and HMT contexts. This notion is consistent with the performance environment of HMTs in much of the literature (e.g., Chen & Barnes, 2012; Gao, Cummings, & Solovey, 2014; Gutzwiller, Lange, Reeder, Morris, & Rodas, 2015). For instance, Chen and Barnes (2012) used a simulation in which a human operator teams with an intelligent agent on a supervisory control task involving multiple robots. The researchers were also attempting to induce a high degree of multitasking.

2.2 Individual Differences

Individual differences is a broad term used to describe characteristics that differ between people. These can include personality traits, intelligence, attitudes, motivation, experience, knowledge, skills, abilities, and other attributes. Individual differences have long been used in personnel selection to identify qualified job candidates and predict performance outcomes (Sackett & Lievens, 2008).

Individual differences have implications for HAI. Findings from multiple studies suggest that the effectiveness of autonomous systems depends on the characteristics of the human operator (e.g., Chen & Barnes, 2012; Lyons & Guznov, 2019; Szalma & Taylor, 2011; Wright, Chen, & Barnes, 2018). Thus, identifying individual differences that account for variability in performance will enhance the effectiveness of autonomous systems via improved personnel selection and job placement.

2.3 Performance Criteria

The goal of this effort is to understand aptitude requirements for jobs involving high levels of HAI. Advances in automated systems will continuously change the performance domain, replacing some tasks performed by the human operator and introducing new tasks and demands. This presents a challenge for researchers, since tethering understanding to task performance with present-day systems will have limited future value.

A recent review article addressing this challenge recommends that the study of human performance with autonomous systems should conceptualize performance as a composite of operator states and distal outcomes (Sibley, Coyne, and Sherwood, 2017). Authors from different disciplines have urged researchers to develop comprehensive mediation models of performance that account for the relationship between stable individual differences, states (e.g., stress), behaviors, and distal outcomes (e.g., Chen et al., 2000; Kanfer, 1990; Zaccaro, 2007).

Developing comprehensive models of performance is especially important for this endeavor since we are concerned with prediction across job types and need to anticipate the impact of future systems. Understanding of the relationships between individual differences, operator states, and performance on different types of tasks will increase the generalizability of findings to different job types and future systems. It is not enough to know *what* is predictive of performance. We also need to understand *why* a trait is predictive, and under what conditions it might become more or less important to performance. For instance, a personality trait might predict performance because of its relationship with stress during multitasking, and might be less relevant for systems and/or job types that have reduced multitasking.

2.4 Research Model for Identifying Predictors of Performance

Our approach to understanding the aptitude requirements of human operators working with autonomous systems accounts for individual differences, operator states, behaviors, task performance, and distal outcomes. This approach will improve understanding of the relevant attributes across different automated systems, and account for boundary conditions (i.e., moderators) of their validity. Refer to Figure 1 for a visual representation of the approach.

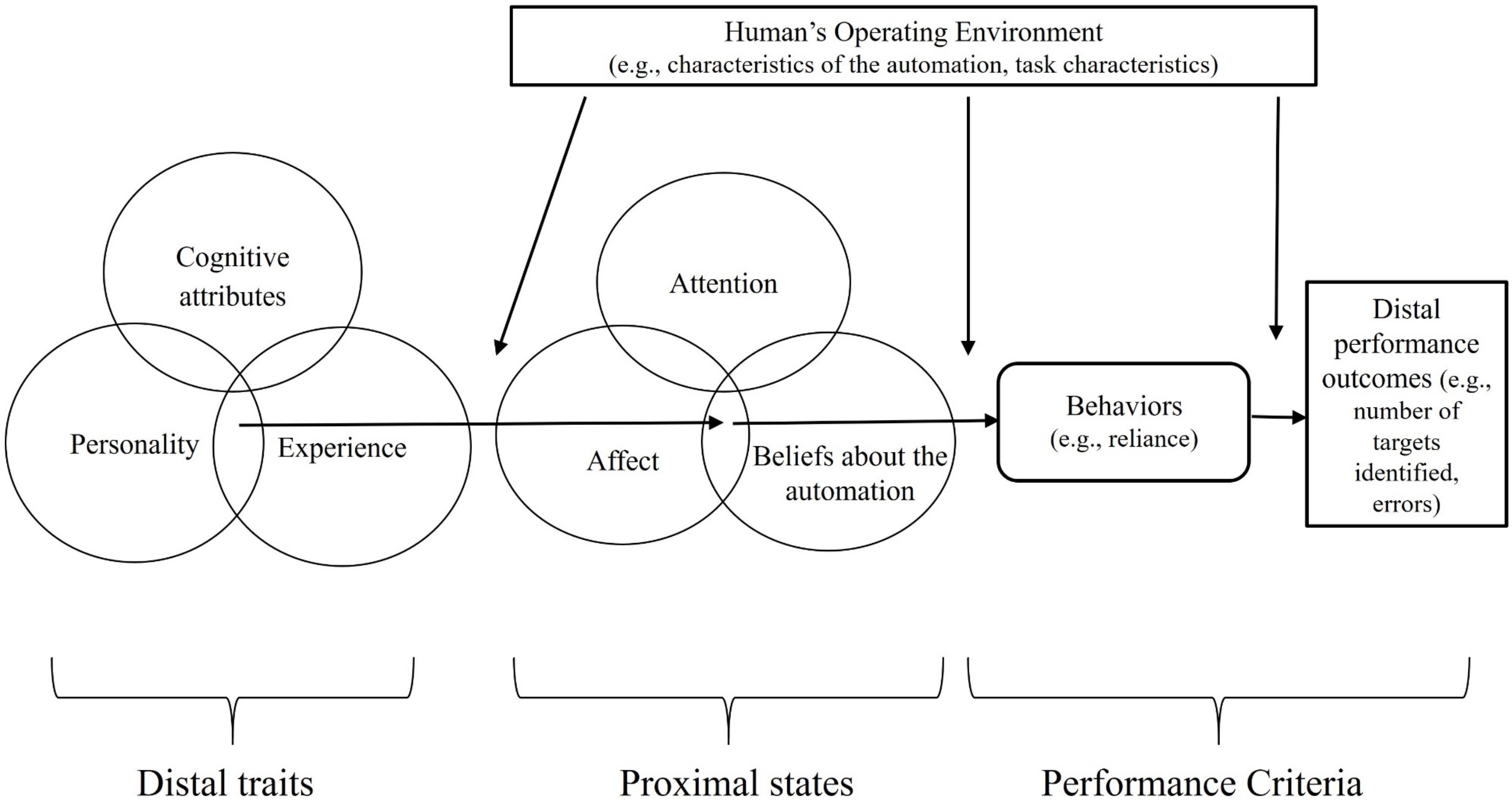

Figure 1. Model of Individual Differences and Performance with Autonomous Systems.

Traits predict operator states and subsequent behaviors, which in turn account for differences in performance indicators. These relationships are moderated by the operating environment.

Note. Adapted from "Leader Traits and Attributes" by Zaccaro, Kemp, and Bader (2004).

3.0 SUMMARY OF RESEARCH GOALS AND OBJECTIVES

This effort had four main objectives. Objective 1 was to identify factors that affect operator aptitude requirements in jobs involving autonomous systems. To accomplish this, we identified task characteristics and operator states relevant to performance.

Objective 2 was to identify entry-level aptitude requirements for jobs involving autonomous systems. Here, we attempted to link the factors identified in Objective 1 to individual differences.

Objective 3 was to determine which entry-level skills, abilities, and other characteristics (SAOC) can be measured by existing DoD tests and where additional assessment methods should be developed. This was partially addressed in the course of objectives 1, 2, and 4.

Objective 4 was to conduct studies to examine the relations between identified critical entry-level SAOCs and performance on tasks requiring interaction with autonomous systems. This was addressed via a meta-analysis of the relationships between individual differences and performance on tasks relevant to HAI.

4.0 OBJECTIVE 1 – FACTORS AFFECTIVING APTITUDE REQUIREMENTS

USAF Airmen jobs are a prime example of job redesign due to increased automation prevalence in the workplace. As more automation has been introduced, job characteristics have changed rather than disappeared (Fahey & Miller, 2017). This section reviews task characteristics of jobs with high levels of HAI and anticipates how these might change with advances in technology. This is followed by a review of operator states relevant to HAI.

4.1 Task Characteristics.

A review of the literature led us to identify common tasks characteristics resulting from increased capabilities and reliability of autonomous systems. The most salient tasks characteristics in the literature were increased monitoring and increased multitasking. Working effectively as part of an HMT is also highly important, especially as autonomous systems become more sophisticated.

4.1.1 Monitoring.

As more tasks are performed by autonomous agents, many Airman jobs will involve fewer manual tasks. Meanwhile, increases in the number of autonomous agents in the performance domain will put greater demands on the ability of human operators to supervise these agents (Chen & Barnes, 2012). Operators will need to maintain their focus of attention and detect infrequent and unpredictable events over extended periods of time (Matthews, Langheim, & Warm, 2011; Reinerman-Jones, Matthews, Langheim, & Warm, 2011; Sheridan, 2002). Future systems are expected to commit fewer errors that are more complex in nature. Increased time between higher-complexity errors will make it more difficult for operators to detect errors, and greater SA will be required to understand the error and react appropriately (Sellner, Hiatt, Simmons, & Singh, 2006).

Performance Criteria for Monitoring Tasks. Objective performance criteria for monitoring tasks in the literature often include reaction time, detection rate (correctly identifying errors made by the automated agent), and false alarms (incorrectly identifying an error when none was made; e.g., Karpinski, Chancey, Palmer, & Yamani, 2018; Molloy & Parasuraman, 1996; Sato, Yamani, Liechty, & Chancey, 2019).

Our review identified a small number of other studies using mind-wandering frequency as a criterion measure for monitoring and sustained attention tasks (e.g., Casner & Schooler, 2015; Gouraud, Delorme, & Berberian, 2018). For instance, Casner and Schooler (2015) used a mind-wandering frequency questionnaire as the performance criterion for airline pilots.

Relevant Operator States. The most relevant operator state during monitoring tasks is attention. Monitoring tasks are also associated with increases in workload, stress, and boredom (Dillard et al., 2019; Helton, Shaw, Warm, Matthews, & Hancock, 2008; Hunter & Eastwood, 2018). Section 5.2 describes these states in more detail.

4.1.2 Multitasking.

Multitasking occurs when a person engages in a variety of tasks simultaneously via task-switching (Chen & Barnes, 2012). As more tasks are delegated to autonomous agents, human operators become responsible for supervising the completion of multiple tasks simultaneously, rather than maintaining their focus on a small number of tasks that they complete manually. Although autonomous agents are expected to reduce the number of manually-completed tasks for the human operator, it is expected to increase the number and variety of tasks for which the human operator is responsible. As a result, operators must continuously switch their attention between multiple streams of information (Richards, Izzetoglu, & Shelton-Rayner, 2017).

Performance Criteria for Multitasking. Indicators of multitasking performance in the automation literature include visual scanning tasks, route editing, and SA (Chen & Barnes, 2012). Studies vary in how they operationalize and measure multitasking performance. Adopting a standard measure of multitasking could also improve the research in this area, since differences among these studies makes it difficult to compare their findings or appropriately include them in a meta-analysis (König, Bühner, & Mürling, 2005).

Tests specifically designed to elicit multitasking might serve as better criteria for examining the role of individual differences. For instance, Poposki et al. (2009) and Barron and Rose (2017) used a computerized multitasking simulation called SynWin, which presents four tasks simultaneously (memory search, arithmetic, visual monitoring, and auditory monitoring). Another popular measure is the Simultaneous capacity/Multitasking scenario (Bratfisch, Hagman, & Puhr, 2002).

Relevant Operator States. The most relevant operator state during multitasking are stress and workload (Chen & Barnes, 2012; Poposki et al., 2009). Section 5.2 describes these states in more detail.

4.1.3 Effective HMT.

The USAF's vision for the future is to develop highly autonomous systems that will work with human operators as part of an HMT. As stated in the DoD's Unmanned Systems Integrated Roadmap, "Robots will evolve from tools into full teammates that are integrated with our soldiers, airmen, marines, and sailors" (Fahey & Miller, 2017, p. 32). Increasingly advanced autonomous agents that are able to interact with human operators in more sophisticated ways will give greater relevance to predictors of effective HMT.

As systems become more autonomous and capable of complex decision-making, it will be increasingly difficult for humans to understand the reasoning process behind the system's output (Chen, 2018). Autonomous agents might perform actions that humans fail to understand. In a given situation, the human operator must decide if it is appropriate to rely on the automation, even when the underlying information and rationale is not fully understood (Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003; Wright, Chen, Barnes, & Hancock, 2017).

Performance Criteria of HMT. Multiple behavioral criteria are used to measure HMT processes. Reliance and compliance behaviors are commonly used as criteria measures to determine how effectively a human operator is working with an autonomous system. Reliance refers to instances in which an operator refrains from a response when the system is silent or indicating normal operation. Compliance refers to instances in which an operator responds when the system issues a signal. Together these behaviors are referred to as dependence (Chancey, Bliss, Yamani, & Handley, 2016). Subjective measures such as a self-report measure of reliance intentions are also used as criteria (Lyons & Guznov, 2019). Reliance and compliance provide useful information but do not capture the full extent of HMT that is envisioned for future systems.

Operator perceptions of autonomous agents are receiving increased attention. Perceptions relevant to an operator treating a technology as a teammate versus a tool include agents' benevolence (Panganiban, Matthews, & Long, 2019), intent (Lasota & Shah, 2015), anthropomorphism (de Visser et al., 2016), and perceived task interdependence (Lyons, Wynne, Mahoney, & Roebke, 2019).

Wynn and Lyons (2018) put forth a construct that incorporates such perceptions, called teammate-likeness, which they define as "the extent to which a human operator perceives and identifies an autonomous, intelligent agent partner as a highly altruistic, benevolent, interdependent, emotive, communicative, and synchronized agent teammate, rather than simply an instrumental tool" (p. 3).

Relevant Operator States. Operator states relevant to HMT include trust (Lee & See, 2004), complacency (analogous to social loafing) and affective reactions such as liking (Merritt, 2011). Section 5.2 describes these states in more detail.

4.2 Operator States Relevant to Performance

Operator states are critical to understanding human performance when working with automation. States have been shown to mediate the relationship between traits and performance, and are potentially more predictive of performance than stable traits (Chen, Whiteman, Gully, & Kilcullen, 2000). Kaber and Endsley (1997) identified potential operator performance issues arising from increased automation. Three of the issues they highlight are highly relevant to operator states. These include failing to detect a problem and appropriately intervene; inappropriate trust in the automation; and loss of SA. Below is a literature review of operator states relevant to performance with autonomous systems.

4.2.1 Trust.

Lee and See (2004) define trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (p. 51). Because of the volume of information demanding the attention of the human operator, they must rely on the automated system to perform some tasks without supervision or input. This requires an appropriate level of trust in the automated system. The DoD recognizes trust as a challenge to the integration of unmanned systems integration in the Unmanned Systems Integration Roadmap: "Manual analysis of raw data is impractical and impossible given the volume, variety, and veracity of the data." Analysts need to be able to trust unmanned systems data analytics and strategies to process, store, fuse, analyze, and report information" (see Unmanned Systems Integration roadmap, (Fahey & Miller, 2017, p. 13).

Trust Calibration. It is important to emphasize that the issue of trust in automation is not a matter of high versus low trust, but appropriate versus inappropriate trust. Trust calibration describes the human operator's evaluation of the automation's reliability relative to their own capability to perform a task or make a decision (Chen et al., 2011). Trust calibration also accounts for situational factors currently affecting the automation and/or the operator (e.g., workload, stress level). For instance, an operator might decide to override the automation's suggestion if one has very high confidence in their abilities on a particular task, even if they believe the automation to be reliable. Conversely, a human operator might think the automation is mediocre, but decide to rely on it if they perceive that a task is beyond their ability, or if they are experiencing a very high workload.

***Trust and Performance*.** Trust has implications for the human operator's effective use of automation, which in turn affects performance on monitoring and multitasking. Having too much trust in automation can result in decreased performance on monitoring tasks due to over-reliance and loss of SA (Parasuraman & Riley, 1997). Having too little trust in automation can result in longer reaction times and missed signals (e.g., from warning systems) due to under-reliance on the automation. Under-reliance can also hinder multitasking performance, and increase the workload of the human operator (Hancock Et Al., 2011).

4.2.2 Complacency.

Weiner (1981, P. 119) Defined Complacency As, "a psychological state characterized by a low index of suspicion." Put simply, complacency is the lack of attention devoted to monitoring automated tasks, paired with an overreliance on automation. When complacent, humans reduce their attention to raw input data that drive the automation and have reduced SA This can be due to excessive trust in the automation (Bagheri & Jamieson, 2004; Metzger & Parasuraman, 2005).

Complacency is more likely to occur when automation is high and when systems are highly reliable (Bagheri & Jamieson, 2004). For This reason, pilots will often reduce the level of automation in an effort to remain vigilant and engaged, especially when task load is low (Bhana, 2010). When required to multi-task, complacency is higher for automated tasks with high reliability—operators shift their attention towards manual tasks and away from automated tasks (Parasuraman & Manzey, 2010).

Complacency and Performance. Complacency is often operationalized as a failure to detect system errors or inadequate response time in detecting those errors. This occurs when a human operator monitoring an automated system engages in suboptimal monitoring behavior, leading to performance failure (Merritt et al., 2019). Multiple studies have linked complacency to task performance outcomes including error detection (Merritt et al., 2019; Wickens, Clegg, Vieane, & Sebok, 2015) and commission errors (following the automation's directive without taking other information into account; Bahner, Hüper, & Manzey, 2008). It is plausible that complacency is a result of inappropriate trust for the automation.

4.2.3 Stress.

Stress refers to a response to external factors known as stressors. Stressors can be direct (e.g., noise), or indirect, such as perceiving that a task exceeds your capabilities (Lin et al., 2015). Increases in workload can elevate stress levels in participants (Panganiban & Matthews, 2014). Tasks requiring continuous monitoring also raise stress (Dillard et al., 2019).

Stress and Performance. Stress causes decrements in cognitive capacities such as working memory and attentional resources. Wohleber et al (2019) found a negative relationship between subjective distress and detection performance in an unmanned ground vehicle (UGV) simulation requiring multitasking. Poposki et al. (2009) also found a negative relationship between stress and multitasking.

4.2.4 Boredom.

Boredom can be defined as an unpleasant affective state in which a person lacks interest in a specific current activity (Fisher, 1993). Boredom is a common challenge among Unmanned Aerial Vehicle (UAV) operators (Button, 2009; Thompson, Lopez, Hickey, DaLuz, & Caldwell, 2006).

Boredom and Performance. Boredom is likely to be induced by long periods of monitoring (Hunter & Eastwood, 2018). However the effects of boredom on performance are difficult to study due to the long period of inaction required to induce boredom. Most studies are not long enough, and/or are trying to measure additional aspects of performance, which requires the inclusion of tasks. Multiple hours of inaction are not practical for many researchers. One study did such an experiment and found higher levels of distraction and worse reaction times when boredom was induced (Cummings et al, 2013).

4.2.5 Situation Awareness (SA)

SA is a term used to describe an operator’s understanding of the relevant elements of the task environment, as well as how these elements might change due to external factors or the operator’s own actions (Loft et al., 2018). SA is likely to remain critical for future airman teaming with highly autonomous teammates. Automated systems are not perfectly reliable, and maintaining SA will allow operators to detect situations in which automation errors are more likely. SA decrements are associated with perceived high system reliability, excessively high levels of trust, and perceived time-pressures (Bahner et al., 2008; Wickens et al., 2016).

SA and Performance. Insufficient SA has contributed to poor human performance when working with autonomous robots (Burke & Murphy, 2004). This is particularly the case for performance on monitoring tasks. A growing problem with highly reliable automation is the tendency to become complacent and fail to monitor agent performance. Multiple studies have found that participants with higher SA perform better than those with lower SA during aviation and naval tasks (Loft, Morrell, & Huf, 2013; Loft et al., 2018). The few studies examining the effect of changes in SA at the within-subjects level found that a change in SA was predictive of change in performance (Endsley, 1990; Loft et al., 2018).

4.2.6 Perceived Workload.

The term *workload* refers to the relationship between task demands and available operator mental capacity (Loft et al., 2018). Workload is typically measured using the National Aeronautics and Space Administration Task Load Index (NASA-TLX) (a self-report measure; Wright et al., 2018). Human operators typically experience variations in workload when assisted by automation. Long periods of monitoring are often punctuated with periods of high perceived workload (e.g., when responding to system errors; Lin et al., 2015).

Perceived Workload and Performance. Multiple studies report a performance decrement on monitoring tasks in response to sudden decreases in workload level (Cumming & Croft, 1973; Goldberg & Stewart, 1980; Matthews, 1986). Cox-Fuenzalida (2000) found the same to be true of sudden workload increases.

5.0 OBJECTIVE 2 – IDENTIFYING APTITUDE REQUIREMENTS

The previous section identified factors affecting human performance in jobs involving HAI. These factors, and their implication for performance, provide the basis of our approach to Objective 2 – identifying aptitude requirements for operators working with autonomous systems. While this is a burgeoning area of study, there are multiple lines of existing evidence suggesting the relevance of certain aptitudes for performance in jobs with autonomous systems. In addition to considering evidence directly from studies of autonomous systems, we reviewed literature on predictors of monitoring and vigilance tasks, as well as predictors of multitasking performance. Such evidence can be extrapolated to HAI, but not to HMT. Evidence pertinent to predictors of HMT is described in separate paragraphs throughout this section.

5.1 Cognitive Skills and Abilities

General mental ability (GMA) is among the strongest and most consistent predictors of job task performance (Schmidt & Hunter, 1998). The criterion-related validity of GMA is generalizable across occupations and job types, with meta-analyses reporting corrected correlations of .40 and above (Schmitt, 2014). Further, the magnitude of the relation between cognitive ability and job performance increases as job complexity increases (Hunter, 1986; Salgado, Anderson, Moscoso, Bertua, de Fruyt, & Rolland). However, it remains important to identify the extent to which specific cognitive attributes contribute to particular job types and outcomes (e.g., working memory) might be more relevant than others for a particular job type. For instance, Hunter and Hunter's (1984) meta-analysis showed that the validity of cognitive ability scores was moderated by job complexity. It is important to examine the relationships between cognitive abilities and performance, in order to understand behavior in a particular domain (Schmitt & Chann, 1998). Below we summarize findings regarding these variables and monitoring, multitasking, and HMT processes.

5.1.1 Working Memory (WM)

WM is a cognitive system that allows for the storage and processing of information needed to execute tasks (Baddeley, 1986; König et al. 2005). WM is typically measured using a combination of memory span tasks, monitoring tasks, and/or switching tasks, such as those in the Oberaur et al. (2003) test battery. These different tasks measure different facets of the WM construct.

HAI. WM has been shown to be closely related to reasoning and GMA (Kyllonen & Christal, 1990; Stauffer, Ree, & Carretta. 1996). Studies have identified WM as important in jobs requiring attentional focus, speed, and accuracy (Paullin, Ingerick, Trippe, & Wasko, 2011; Waters, Russell, Shaw, Allen, Sellman, & Geimer, 2000). WM has been found to be predictive of task performance when working with automation in a simulator (Ahmed et al., 2014; de Visser, Shaw, Mohamed-Ameen, & Parasuraman, 2010; McKendrick et al., 2014; Saqer & Parasuraman, 2017).

There is evidence that individuals with greater WM are better at maintaining alertness during long periods of inaction. This has been attributed to differences in intrinsic alertness in the absence of external cues (Unsworth et al., 2009; Unsworth & Robinson, 2020).

Studies have found WM to be the best predictor of multitasking performance, followed by reasoning and attention (Bühner, König, Pick, & Krumm, 2006; König et al. 2005). Previous research has recommended using WM measures in personnel selection for jobs with multitasking requirements (Colom, Martinez-Molina, Shih, & Santacreu, 2010).

HMT. WM might be relevant to human-machine team processes to the extent that human operators must hold multiple pieces of information received from autonomous teammates for effective decision making and collaboration. A similar line of reasoning has been used in the human-human teaming literature (Mojzisch, Krumm, & Schultze, 2014).

5.1.2 Spatial Ability (SpA).

SpA represents a person's ability to perform complex mental transformations (e.g., spatial visualization; Fincannon, Keebler, Jentsch, & Curtis 2013). Measures of SpA most common in the literature reviewed include the Cube Comparison Test (Ekstrom et al., 1976), and the Spatial Orientation Test (modeled after the Cardinal Direction Test developed by Gugerty & Brooks, 2004). It should be noted that there is some ambiguity in the literature, with SpA being used broadly to describe a variety of visuospatial constructs. It is therefore important to note the measure used when comparing findings across studies.

HAI. SpA has been linked to performance on tasks requiring visual scanning (e.g., target detection, route editing; Chen & Barnes, 2012). It also has been found to be important for multitasking and monitoring (Morgan et al., 2011).

HMT. The relationship between spatial ability and HMT processes is less clear. These variables might not be as closely related if there is a not a visual element to the performance domain, which is the case with most studies using simulators.

5.1.3 Fluid Intelligence.

Fluid intelligence is the ability to reason and to solve novel problems. Measures of fluid intelligence most common in the literature reviewed include the Letter Sets Test (Ekstrom, French, Harmon, & Derman, 1976), the Raven Advanced Progressive Matrices (RAPM); (Raven, Raven, & Court, 1998), and the Intelligence Structure Test 2000 R (Amthauer, Brocke, Liepmann, & Beauducel, 2001).

HAI. Fluid intelligence has been found to be predictive of monitoring tasks (Matthews et al., 2007; Schweizer & Moosbrugger, 2004; Shaw et al., 2010) and multitasking performance (Bühner et al., 2006; König et al., 2005). This is likely attributable to the amount of cognitive resources drawn upon by these tasks. There is some evidence that WM mediates the relationship between fluid intelligence and performance (Held & Carretta, 2013).

HMT. Regarding HMT processes, fluid intelligence is likely to be especially relevant to the operator's strategic use of the autonomous system. Several studies have shown that the relation of cognitive ability to job performance increases as job complexity increases (e.g., Hunter, 1986; Salgado et al., 2004). Operators with greater reasoning ability might be better at identifying the strengths and weaknesses of an automated system, and adapt their strategy accordingly.

5.1.4 Attentional Control (AC).

AC refers to an individual's ability to flexibly shift and focus their attention between different aspects of the environment, allowing individuals to maintain useful information and disengage from unimportant information (Derryberry & Reed, 2002; Foroughi et al. 2019; Thropp et al., 2018). The AC Scale (Derryberry & Reed, 2001) is the most commonly used measure. It is a self-report scale, measuring attention focus and attention shifting ability. Another related measure is the cognitive failures questionnaire (CFQ; McVay & Kane, 2009).

HAI. Multiple lines of evidence suggest a positive relationship between AC and HAI. A recent study found a positive relationship between AC and monitoring performance during a simulated multiple UAV mission (Levulis, DeLucia, & Kim, 2018). There is also evidence suggesting that people higher on AC can devote attentional resources to a greater number of tasks without sacrificing performance (Chen & Barnes, 2012; Chen & Joyner, 2009).

HMT. Chen and Barnes (2014) identified AC as potentially having relevance to HMT. Indirect evidence is seen in studies finding that AC has a positive correlation with SA, and negative correlations with complacency and perceived workload (Chen & Barnes, 2012; Chen & Joyner 2009; Chen & Terrence, 2009; Levulis et al., 2018; Wright et al., 2018). Greater SA among those high on AC could allow for better awareness and understanding of the autonomous agent (e.g., its strengths and weaknesses under various conditions), leading to outcomes such as accurate trust calibration. The relationships with complacency and workload might suggest that those higher on AC are less likely to over-rely on the autonomous agent.

5.2 Big Five Personality Traits

5.2.1 Conscientiousness.

Conscientiousness is the quality of being self-disciplined, careful, reliable, and organized (Duckworth, Peterson, Matthews, & Kelly, 2007). Conscientiousness is usually measured using a 10-item or 20-item self-report measure using a Likert-type scale (Goldberg et al., 2006).

HAI. Conscientiousness has been found to predict job performance in a variety of domains (Barrick, Mount, & Judge, 2001), and for vigilance and attention tasks in particular (Rose, Murphy, Byard, & Nikzad, 2002) as well as task-engagement during performance (Matthews et al., 2007). Finomore et al. (2009) noted that the carefulness and perseverance facets of conscientiousness might promote vigilance, while the impulse-control aspect promotes sustaining attention. Thus, high scores on conscientiousness measures will be positively related to performance in situations requiring vigilance and attention.

HMT. Regarding HMT, the achievement striving and perseverance facets of conscientiousness would likely drive a positive relationship. Although one study indicates that conscientiousness might be associated with under-reliance on autonomous systems (Lin, 2017), such an effect could likely be due to naiveté of the capabilities of the autonomous teammate. Thus, training and experience might be a boundary condition to this relationship (since these would allow someone to improve their performance). In human-human teams, a moderate corrected relationship between conscientiousness and teamwork has been observed (Barrick et al., 2001).

5.2.2 Agreeableness.

Agreeableness is characterized by qualities such as cooperation, trustfulness, compliance, and affability (Barrick et al., 2001). Agreeableness is usually measured using a 10-item or 20-item self-report measure using a Likert-type scale (Goldberg et al., 2006).

***HAI*.** There is not a compelling empirical or theoretical link between agreeableness and monitoring and multitasking. Studies have found weak relationships to HAI outcomes (Lin, 2017; Robinson, Miller, & Unsworth, 2020).

HMT. Among the Big Five personality traits, agreeableness has the strongest correlation with teamwork across professions and job types (Barrick et al., 2001). Trust is a facet of agreeableness that is particularly relevant to performance in this context. *In* human-to-human teams, intra-team trust has been linked to team performance, especially for tasks with higher interdependence between teammates (De Jong, Dirks, & Gillespi, 2016). Trust development will likely be different in human-autonomous agent teams, as the human might not always understand the information and decision-making processes used by the agent, less rich communication media, and different affective responses. Developing the trust required for effective team performance will be especially difficult in these circumstances. Similar to what has been found for virtual teams (Breuer et al., 2016), a high level of dispositional trust is likely to be required for successful performance in this context. Indeed, Huang and Bashir (2017) found dispositional trust was positively related to participants' initial trust for an automated agent. They suggest that people with higher dispositional trust in humans might be more likely to see an automated agent as human, resulting in greater trust.

5.2.3 Emotional Stability.

Emotional stability is characterized by lack of anxiety, hostility, depression, and personal insecurity (Barrick et al., 2001). Emotional stability is usually measured using a 10-item or 20-item self-report measure using a Likert-type scale (Goldberg et al., 2006).

***HAI*.** Previous research supports a mediation model in which neuroticism is negatively related to multitasking performance via its relationship with state anxiety (Popowski, Oswald, & Chen, 2009). Trait anxiety is a facet of neuroticism relevant to performance with autonomous systems. State anxiety when measured via physiological response has been found to reduce the accuracy of spatial ability of working memory (Shackmann et al., 2006).

HMT. When working with automated agents, positive attitudes and emotions have been found to foster trust, reliance, and satisfaction, whereas negative emotions lead to disuse (Schaeffer et al., 2016). It has been suggested that a person's decision to rely on an automated system is driven by emotion as much as it is driven by reasoning (Lee & See, 2004). It has been found that affective states (e.g., happiness) during a human-agent interaction are predictive of trust, liking, and reliance on automated systems (Merritt, 2011). Many features of the automated agent are designed to improve the user's affective reactions and subsequent satisfaction, however people differ in their propensity for positive or negative emotions. Such individual differences between users are a strong determinant of affective reactions to automation, and are therefore predictive of user satisfaction during a human-agent interaction (Merritt, 2011; Merritt, Heimbaugh, LaChapell, & Lee, 2013; Schaefer, Chen, Szalma, & Hancock, 2016). A moderate relationship has been found between emotional stability and teamwork and human-human teams (Barrick et al., 2001).

5.2.4 Extraversion.

Extraversion refers to the extent to which someone is outgoing and energetic versus solitary and reserved (Eysenck, 1983). Extraversion is usually measured using a 10-item or 20-item self-report measure using a Likert-type scale (Goldberg et al., 2006).

HAI. There is some support for the notion that people high on extraversion have greater capacity on cognitive skills such as working memory and divided attention, and that they tend to do better on tasks involving multitasking, but worse on tasks requiring sustained attention (Matthews, 2003; Unsworth et al., 2009). Findings for multitasking are conflicting between studies however (e.g., König et al., 2005; Poposki et al., 2009).

HMT. The relevance of extraversion to HMT is ambiguous. One could argue that high extraversion would result in better teaming via processes such as collaboration, engagement, and knowledge sharing. However, the evidence from human-human teams suggests a weak relationship (Barrick et al., 2001).

5.2.5 Openness to Experience.

Openness to experience describes how willing a person is to accept new ideas, activities, and routines. People high on openness are more likely to describe themselves as reflective, introspective, and curious. They prefer novelty, variety, and complexity. People low on openness are more likely to prefer convention, routine, simplicity, and utilitarianism (Rose, Barron, Carretta, Arnold, & Howse, 2014). While openness to experience has a weak relationship with overall job performance across occupations, it can be useful for predicating specific criteria (e.g., training performance; Barrick et al. 2001). Openness is usually measured using a 10-item or 20-item self-report measure using a Likert-type scale (Goldberg et al., 2006).

HAI. Overall, there is not compelling evidence for the efficacy of openness as a predictor of HAI. Rose et al. (2014) found a negative relationship with remotely piloted aircraft (RPA) training performance. This was attributed to differences in decisiveness during performance, which was deemed critical for the job. Unsworth et al. (2009) found no relationship between openness and vigilance. Another study found similar results for performance in a simulator with highly automated aviation tasks (Eschen et al., 2016). Specifically, the authors found longer reaction times and lower accuracy for participants higher on a specific facet of openness known as openness to actions. This finding might be due to a preference for reflection among those high on openness, which could interfere with sustaining vigilance and multitasking.

HMT. Given the weak relationship with performance in general, it is unlikely that openness would be an effective predictor of HMT. Barrick et al.'s (2001) meta-analysis indicates it has the weakest relationship with teamwork in human-human teams. However, the intellectual curiosity facet might have some theoretical relevance to differences in people's willingness to collaborate with autonomous agents and develop an understanding of agent's strengths and weaknesses.

5.3 Other Characteristics

5.3.1 Boredom Proneness (BP).

Bhana (2009) describes BP as, "…a trait referring the propensity of an individual to become bored." BP is often measured using the 28-item BP Scale (Farmer & Sundberg, 1986).

HAI. Findings show that BP is correlated with vigilance, attention, and performance outcomes (Cummings et al., 2016). According to Hunter and Eastwood (2018), trait boredom is highly correlated with failures of sustained attention.

HMT. The relevance of BP for HMT is less clear. One could argue that those high on BP might be prone to decreased SA, which would limit their ability to calibrate trust for and over-reliance on autonomous agents they are teaming with. In addition, those high on BP might attempt to reduce boredom by increasing their workload and under-relying on autonomous agent teammates.

5.3.2 Perfect Automation Schema (PAS).

PAS describes a cognitive schema containing beliefs of how automated systems operate (Lyons et al., 2016; Merritt, Unnerstall, Lee, & Huber, 2015). PAS involves two factors: high expectations (HE) and All-or-None (AoN) thinking. HE reflects beliefs about the reliability of automated systems (e.g., "Automated systems have 100% perfect performance;" Merritt et al., 2015). AoN reflects one's tendency to view automated systems as either perfectly functional or completely useless (e.g., "If an automated system makes an error, then it is broken;" Merritt et al., 2015). The factors are related but show differential prediction of outcomes. Merritt et al. (2015). PAS scale is the most commonly-used measure of PAS.

HAI. Although PAS likely predicts HAI outcomes such as vigilance and multitasking (Merritt et al., 2019), it has greater relevance and potential for predicting HMT processes and outcomes.

HMT. PAS has greater relevance to HMT than HAI. Although the construct has only recently received attention, existing evidence suggests PAS is highly relevant to HMT. High AoN predicts greater initial trust, but steeper declines in trust when automation errors occur (Merritt et al., 2015). Another study found HE, but not AoN, was predictive of fighter pilots' trust in an automatic ground collision avoidance technology (Lyons & Guznov, 2019).

5.3.3 Coping Style.

Coping style refers to one's strategy to handle stressful situations (Matthews & Campbell, 2009). The three types of coping styles are task-focused, emotion-focused, and avoidance. Task-focused coping involves efforts aimed at solving the problem causing stress. Emotion-focused coping strategies attempt to directly deal with the emotions being experienced (e.g., by releasing pent-up frustration; focusing on the positive aspects of the situation). Avoidance coping involves trying to ignore, forget, or remain in denial about a source of stress (Endler & Parker, 1990). Coping is commonly measured using the Coping Inventory for Task Situations (CITS); (Matthews & Campbell, 1998).

HAI. Coping style has received attention among HAI researchers due to its relevance to stress during performance. Moderate correlations have been found between coping style and performance on monitoring tasks, as well as vigilance in a battlefield environment (Matthews et al., 2007; Matthews et al., 2014; Shaw et al., 2010).

HMT. The relevance of coping style to HMT processes and outcomes is less clear and has not received attention in the literature. It could have implications for how people alter their approach to teaming with autonomous agents in response to stressful situations. Coping style could also have implications for perceptions of autonomous agent teammates (e.g., people with emotion-focused coping might attribute performance issues to the autonomous agent rather than consider strategies to work more effectively with the agent to improve performance outcomes.

5.3.4 Video Game Experience (VGE).

VGE is usually operationalized as the frequency and duration of video game play over a recent period (e.g., the past six months). Thus, it is an indicator of proclivity rather than skill at video games. VGE measures also consider genre most often played. However, genre labels are ambiguous (e.g., "action," "racing," "shooter," "multiplayer"), and rely on participants to discriminate between these labels. Genres are also used inconsistently across studies.

HAI**.** VGE has been shown to be predictive of task performance when using autonomous systems (McKinley et al., 2011; Lin et al., 2015). The reason for the relationship between VGE and performance with automated systems is not well-understood, but has important implications for whether interest in video game reflect a stable underlying trait or a consequence of interactions with them (i.e., enhance skills and abilities).

HMT. While the relationship with HAI is well-studied, there is relatively little empirical research linking VGE to HMT states, processes, and outcomes, however it potentially has relevance to HMT processes and outcomes. VGE has been theoretically linked to relevant HMT variables such as trust calibration, and attitudes towards robots (Correia, Mascarenhas, Prada, Melo, & Paiva, 2018; Wang, Pynadath, & Hill, 2016).

Underlying Constructs of VGE**.** Claims of a causal relationship in which video games improve cognitive skills are common in the literature, to the point that it has nearly become conventional wisdom. The publication most widely-cited as evidence of this is very misleading (e.g., it is titled "Action video game modifies visual selective attention;"; Green & Bavelier, 2003). A close study of the paper reveals that the causal relationship suggested by its title and abstract is based on quasi-experiments comparing extreme groups (e.g., heavy users of video games compared vs. people who have never played video games).

Not surprisingly, other authors have consistently failed to replicate these causal results. A recent meta-analysis (Sala, Tatlidil, & Gobet, 2018) examined the correlation between video game *skill* (note: skill in video games is different than hours played) and cognitive ability, as well as the effects of video game training on participants' cognitive ability. It found low correlations between video game skill and cognitive abilities, and no evidence of a *causal* relationship between playing video games and changes in cognitive ability (Sala et al., 2018).

Study design weaknesses and opportunities to improve this area of understanding has only recently been discussed (e.g., Salas et al., 2017; Simons et al., 2016; Waris et al., 2019). Unsworth et al. (2015) argued that the extreme groups approach used in many study designs has resulted in overestimated effects of video games on cognitive ability. This conclusion not been without controversy (e.g., Green et al., 2017 directly responded to this argument). Thus, there is a clear need to explore measurement methods in this area.

6.0 QUANTITATIVE REVIEW: META-ANALYSIS OF EXISTING RESEARCH

After a qualitative review of the factors likely to affect operator aptitude requirements in jobs involving a high level of HAI, it was notable that the extant knowledge of individual differences and performance when working automation are from human factors experiments with small sample sizes, and findings varied between studies. To address this, we conducted a meta-analysis of the extant literature in order to empirically estimate these relationships, and identify areas of future research.

6.1 Research Questions

In addition to the relationships between individual differences and performance with automation, we also wanted to estimate the relationships between individual differences and performance on tasks similar to those performed by Airmen in jobs involving autonomous systems. Studies examining the relationship between individual differences and HMT processes were not plentiful enough to include in the meta-analysis.

We also wanted to estimate relationships between individual differences and important operator states, to the extent that such data was available. Thus, the following research questions were explored:

6.1.1 Research Question 1.

What individual differences are most strongly related to performance when assisted by automation/autonomous agents?

6.1.2 Research Question 2.

What individual differences are most strongly related to performance on monitoring tasks (e.g., vigilance tests)?

6.1.3 Research Question 3.

What individual differences are most strongly related to measures of multitasking ability?

6.1.4 Research Question 4.

What individual differences are most strongly related to operator states of stress, trust, and perceptions of autonomous systems?

6.2 Method

6.2.1 Inclusion Criteria.

Our inclusion criteria were modeled after previous meta-analyses (e.g., Schaefer et al., 2016). To be eligible for inclusion in our data analyses, studies had to meet the criteria described below.

Inclusion Criteria 1. Studies must measure one or more of the individual differences and outcomes of interest.

Inclusion Criteria 2. For studies coded as task performance with automation, participants needed to interact or be assisted by autonomous systems during task performance. While many studies focused on technologies that are associated with automation (e.g., UAV operation), we carefully reviewed the methods to ensure this was the case. For instance, we excluded correlations from studies in which participants tele-operate a UAV (analogous to driving a remote controlled car) without assistance from automation.

6.2.2 Literature Search.

Studies were located using ProQuest, PsycInfo, and google scholar. We used the following search terms: automation, Big Five, personality, individual differences, stress, boredom, vigilance, SA. Additional studies were found in the references sections of related works. The limited amount of research done in this area limited the constructs we were able to examine. Studies were found for all Big Five traits and cognitive attributes. Other variables that were prevalent in the literature (and therefore had ample data available) were included as well.

Since many studies published in this area were not focused on the relationship between personality and performance on their tasks, they did not report correlations. In some cases, statistics were reported from which the correlations could be derived; in other cases we successfully obtained the necessary data or unreported correlations from the authors.

6.2.3 Data Analysis.

Analyses were conducted using r statistical software using formulas taken from Hunter and Schmidt (2004). Composite correlations were computed for studies that reported multiple non-independent effect sizes. We used a random effects model since it is assumed that studies used in the analyses will differ in sample characteristics, methods, etc.

6.3 Results for HAI Task Performance in (Research Question 1)

6.3.1 Cognitive Variables.

Not surprisingly, cognitive variables had the strongest relationships with performance when working with automation. Spatial ability had the strongest relationship ($r^{sw} = .52$, $p < .01$; r^{sw} denotes a sample-weighted correlation) followed by working memory ($r^{sw} = .43$, $p < .01$). A moderate relationship was found between AC and performance ($r^{sw} = .33$, $p < .01$). AC was included as a cognitive variable, however it is measured via self-report scales. Studies could not be found for general mental ability, fluid intelligence, or crystallized intelligence. Results are presented in Table 1.

Table 1. Results for Individual Differences and HAI Performance

				$r^{sw}CI_{95}$	
Predictor	***k***	***N***	r^{sw}	**LL**	**UL**
Cognitive variables					
Spatial Ability	10	340	.52**	.81	.74
Working Memory	4	156	.43**	.29	.55
Attentional Control	2	112	.33**	.15	.79
Big Five factors					
Extraversion	2	165	-.04	-.44	.36
Conscientiousness	2	121	.26**	.08	.42
Neuroticism	2	315	-.17**	-.28	-.06
Agreeableness	2	250	.15*	.02	.27
Openness	1	154	-.07	-.31	.17
Other characteristics					
Boredom Proneness	6	573	.31**	.23	.38
Video Game Experience	4	261	.30**	.18	.41

Note. r^{sw} = sample-weighted correlation; $r^{sw}CI_{95}$ = 95 percent confidence interval (CI); lower limit (LL) of the CI; upper limit (UL) of the CI; *result is significant at the .05 level; **result is significant at the .01 level

6.3.2 Personality

Given the small number of studies that have examined the Big Five and performance with automation, we emphasize that these results should be viewed as a quantitative summary of the extant studies in this nascent area, and used to inform future research. This is evidenced by the wide confidence intervals (CI) associated with these estimates (see Table 1).

The relationships with performance were strongest for conscientiousness and neuroticism (respectively: $r^{sw} = .26$, $p < .01$; $r^{sw} = -.17$, $p < .01$). Conscientiousness had an especially wide confidence interval (.08 to .42). The sample-weighted correlation estimate for extraversion was near-zero, however it had an extremely wide-confidence interval (-.44 to .36). The few studies conducted in this area prevented us from testing for moderation.

Agreeableness had a stronger relationship than many would expect ($r^{sw} = .15$, $p < .05$). Openness to experience had a small and non-significant relationship to performance with automation.

6.3.3 Other Characteristics.

BP and VGE both had positive and significant sample-weighted correlations with performance (respectively: $r^{sw} = .31$, $p < .01$; $r^{sw} = .30$, $p < .01$). Studies examining these characteristics were much more prevalent than for the Big Five personality traits.

6.4 Results for Monitoring Performance (Research Question 2)

6.4.1 Cognitive Variables.

Fluid intelligence had the strongest relationship with measures of multitasking ability (r^{sw} = .40, p < .01). The only other cognitive variable approaching this strength was working memory (r^{sw} = .31, p < .01). AC had a moderate to strong relationship with multitasking (r^{sw} = .28). See Table 2.

Table 2. Results for Individual Differences and Monitoring

				$r^{sw}CI_{95}$	
Predictor	***k***	***N***	**r^{sw}**	**LL**	**UL**
Cognitive variables					
Working Memory	3	472	.31**	.22	.39
Fluid Intelligence	4	899	.40**	.28	.51
Crystalized Intelligence	2	672	.18**	.06	.29
Attentional Control	3	796	.16**	.09	.23
Big Five factors					
Extraversion	4	757	-.15**	-.22	-.08
Conscientiousness	2	440	.19**	.10	.28
Neuroticism	3	547	-.15**	-.29	.00
Agreeableness	3	547	.12*	-.04	.27
Openness	2	440	.07	-.02	.16
Other characteristics					
Emotion-focused coping	3	779	-.16**	-.24	-.07
Task-focused coping	3	779	.34**	.27	.40
Avoidance coping	3	779	-.33**	-.44	-.21

Note. r^{sw} = sample-weighted correlation; $r^{sw}CI_{95}$ = 95 percent confidence interval (CI); lower limit (LL) of the confidence interval; upper limit (UL) of the CI;*result is significant at the .05 level; **result is significant at the .01 level

6.4.2 Personality.

Extraversion has received more attention in the monitoring and vigilance literature than other personality traits. We found a negative relationship between extraversion and monitoring (r^{sw} = -.15, p < .01). Similar results were found for neuroticism (r^{sw} = - .15, p < .01). The strongest relationship was between conscientiousness and performance (r^{sw} = .19, p < .01). Sufficient studies were located for agreeablenesss and openness to experience, however their relationships with monitoring performance did not reach significance.

6.4.3 Other Characteristics.

Our literature search identified multiple studies that examined coping style. Analyses revealed strong correlations between these self-report measures and performance on monitoring tasks. Task-focused coping and avoidance coping had the highest relationships with performance (respectively: $r^{sw} = .34$, $p < .01$; and $r^{sw} = -.33$, $p < .01$).

6.5 Results for Multitasking Performance (Research Question 3)

6.5.1 Cognitive Variables.

Fluid intelligence had the strongest relationship with measures of multitasking ($r^{sw} = .42$, $p < .01$). WM had a moderate to strong relationship ($r^{sw} = .27$, $p < .01$). AC had a small but significant correlation with multitasking ($r^{sw} = .16$). Results are presented in Table 3.

6.5.2 Personality.

Studies examining personality in the context of multitasking were relatively numerous. However, the correlations were much smaller than was found for vigilance. Most surprisingly, conscientiousness had a negative and significant relationship to multitasking ($r^{sw} = -.07$, $p < .01$). Every study reported a negative correlation. This is in line with much of the multitasking literature and gives credence to theories in the HAI literature suggesting that the attention to detail facet of conscientiousness could hinder performance.

Neuroticism had a similar relationship as it did for monitoring ($r^{sw} = -.16$, $p < .01$). Most studies attributed this to the stress and anxiety facets of neuroticism.

6.5.3 Other Characteristics.

Stress is a common theme in the multitasking literature. Some studies examined stress tolerance instead of global neuroticism. However, our analysis of these studies found a near-zero relationship ($r^{sw} = .03$, $p > .05$). Polychronicity, a term describing one’s preference for multitasking, is also common in this literature. However we found a weak relationship for this trait ($r^{sw} = .06$, $p > .05$).

Table 3. Results for Individual Differences and Multitasking

				$r^{sw}CI_{95}$	
Predictor	***k***	***N***	r^{sw}	**LL**	**UL**
Cognitive variables					
Working Memory	3	345	.27	.17	.37
Fluid Intelligence	6	3,339	.42	.32	.51
Attentional Control	2	243	.28	.16	.39
Big Five factors					
Extraversion	5	3,217	.07	.07	.03
Conscientiousness	4	3,096	-.07	-.11	-.03
Neuroticism	4	546	-.16	-.24	-.07
Agreeableness	3	2,943	.00	-.03	.04
Openness	4	3,096	.03	-.09	.15
Other characteristics					
Stress Tolerance	2	2,861	.03	.00	.07
Polychronicity	2	2,861	.06	-.03	.15

Note. r^{sw} = sample-weighted correlation; $r^{sw}CI_{95}$ = 95 percent confidence interval (CI); lower limit (LL) of the confidence interval; upper limit (UL) of the CI;*result is significant at the .05 level; **result is significant at the .01 level

6.6 Results for Individual Differences and Operator States

Our literature search found sufficient studies to estimate relationships for the Big Five and stress, the Big Five and trust, and for PAS and trust. Finally, we report results for PAS and perceptions of automated systems. We were unable to estimate relationships for workload, complacency, or SA, due to a lack of a sufficient number of studies with outcomes in common.

6.6.1 Individual Differences and Stress.

A sufficient number of studies linking the Big Five to stress were found. Neuroticism had the only positive relationship with stress. This was also by far the strongest relationship with stress. Extraversion and conscientiousness both had significant but modest relationships with stress. Significant relationships were not found for agreeableness or openness to experience. See Table 4.

Table 4. Results for Individual Differences and Stress

				$r^{sw}CI_{95}$	
Predictor	***k***	***N***	r^{sw}	**LL**	**UL**
Big Five factors					
Extraversion	4	4,583	-.14**	-.23	-.05
Conscientiousness	4	4,644	-.15*	-.26	.03
Neuroticism	4	4,692	.38**	.27	.48
Agreeableness	3	4,434	-.11	-.25	.03
Openness	3	4,434	-.05	-.16	.06

Note. r^{sw} = sample-weighted correlation; $r^{sw}CI_{95}$ = 95 percent confidence interval (CI); lower limit (LL) of the confidence interval; upper limit (UL) of the CI;*result is significant at the .05 level; **result is significant at the .01 level

6.6.2 Individual Differences and Trust.

Because two meta-analyses have been conducted on Individual differences and trust (Hancock et al., 2011; Schaeffer et al., 2016), we only were interested in individual differences not covered by these meta-analyses. Many of the studies found used behavioral measures of trust (e.g., trust games). For the sake of consistency and clarity of the construct, we only included studies using self-report measures of trust. Detailed results are presented in Table 5.

Among the Big Five, moderate and significant relationships were found for agreeableness and conscientiousness. Neuroticism had a lower correlation than many would expect.

The PAS factors differed greatly in their relationship to trust. The expectations facet had a moderate and significant relationship with trust, while the all or nothing facet had a near-zero relationship. It should be noted that these estimates come from studies conducted by the same authors, which could bias findings.

Table 5. Results for Individual Differences and Trust

				$r^{sw}CI_{95}$	
Predictor	***k***	***N***	**r^{sw}**	**LL**	**UL**
Big Five factors					
Extraversion	2	447	.12	-.10	.32
Conscientiousness	2	447	.19*	.03	.34
Neuroticism	2	447	-.12	-.41	.19
Agreeableness	2	447	.27**	.18	.36
Openness	2	447	.09	-.21	.36
Perfect Automation Schema					
PAS All or Nothing	5	492	-.05	-.25	.15
PAS Expectations	6	634	.24**	.13	.34

Note. r^{sw} = sample-weighted correlation; $r^{sw}CI_{95}$ = 95 percent confidence interval (CI); lower limit (LL) of the confidence interval; upper limit (UL) of the CI;*result is significant at the .05 level; **result is significant at the .01 level

6.6.3 PAS and Positive Perceptions.

For PAS studies, we collapsed outcomes related to operator evaluations of the automation into a broader outcome we labeled “positive perception.” The outcomes collapsed into positive perception were liking, benevolence, reliability perceptions, and performance perceptions. Collapsing these outcomes allowed us to estimate relationships for overall positive judgments of automation. See Table 6.

It should be noted that these estimates are based on correlations from three studies published in the same article, and all correlations come from the same lead author.

Table 6. Results for PAS and Positive Perceptions.

				$r^{sw}CI_{95}$	
Predictor	***k***	***N***	**r^{sw}**	**LL**	**UL**
Perfect Automation Schema					
PAS All or Nothing	3	268	-.05	-.17	.07
PAS Expectations	3	268	.18*	.01	.35

Note. r^{sw} = sample-weighted correlation; $r^{sw}CI_{95}$ = 95 percent confidence interval (CI); lower limit (LL) of the confidence interval; upper limit (UL) of the CI;*result is significant at the .05 level; **result is significant at the .01 level

6.7 Discussion

6.7.1 Cognitive Skills and Abilities.

Cognitive attributes were strong predictors of performance. Understanding the specific cognitive skills and abilities that contribute to performance can shed light on why some variables such as VGE are predictive of performance in jobs with high levels of automation. WM might be especially useful as a selection tool, and tests likely exist in DoD test batteries. Previous research has recommended using WM measures in personnel selection for jobs with multitasking requirements (Colom, Martinez-Molina, Shih, & Santacreu, 2010).

6.7.2 Big Five Personality Traits.

Among the Big Five, conscientiousness and neuroticism had the strongest correlations with outcomes.

Conscientiousness. The large confidence interval found for the relationship between conscientious and performance with automation was surprising. The negative relationship found for multitasking suggests a moderating role of task characteristics. A review of the literature identified characteristics of high-automation jobs that likely have implications for this relationship.

While conscientiousness is likely beneficial to vigilance and attention tasks (Rose, Murphy, Byard, & Nikzad, 2002), this is not necessarily the case for multitasking. Along with this concern, there is also the notion that high conscientiousness individuals are prone to under-rely on automation in exchange for personally completing the tasks and being aware of details (e.g., the information used by the automation to reach a decision). When workload is very high and automation is accurate, this can be detrimental to performance. Indeed, Lin (2017) found conscientiousness was the only Big Five trait negatively related to reliance on automation during high workload conditions; and conscientiousness had a negative relationship with performance (worse accuracy, more missed tasks). Performance detriments associated with conscientiousness could likely be mitigated with training and an improved understanding of how the automation can aid performance.

A chief goal in the field of automation is to develop automated aids that complement human abilities. Since automation will be more capable at some tasks than humans (Musić & Hirche, 2017), human operators will need to relinquish control at times to achieve optimal performance. The positive relationship found for trust suggests that conscientiousness might not hinder this process if the human operator has enough information to understand the optimal way to work with the automation (e.g., by not interfering with excessive attention to detail at the expense of completing other tasks).

Neuroticism. Results indicate that neuroticism might be predictive of performance for jobs with high levels of automation. Reasons for this finding are likely due to neuroticism’s relationship with affect, and its relationship to deficits in working memory and multitasking. When working with automated agents, positive attitudes and emotions have been found to foster trust, reliance, and satisfaction, whereas negative emotions lead to disuse (Schaefer et al., 2016). Specifically, the facets of vulnerability and anxiety reflect one’s tendency to experience negative emotions such as stress and ambivalence.

Other explanations for the importance of neuroticism focus on the cognitive demands of jobs with high levels of automation. Multitasking and memory components of monitoring tasks. Previous research supports a mediation model in which neuroticism is negatively related to multitasking performance via its relationship with state anxiety (Poposki, Oswald, & Chen, 2009). Furthermore, state anxiety (when measured via physiological response) has been found to reduce the accuracy of spatial or working memory (Shackman et al., 2006). Our findings for the relationship between neuroticism and stress lend credence to this explanation.

Extraversion. The small correlation found for extraversion and automation performance is in line with the traditional understanding of this trait. However, there was a wide confidence interval. The opposing results for monitoring and multitasking suggest that the relationship with automation could depend on the task type. Unfortunately, there were not enough studies to test for moderation. There have been theories posited in the literature that extraversion might positively predict monitoring and multitasking, however this effect is not reflected in our findings. The low relationships with trust and stress are consistent with the rest of our findings.

Agreeableness. The relationship found between agreeableness and performance with automation led us to explore possible explanations. We also found a relationship between agreeableness and trust, which is believed to be important for performance on tasks requiring high levels of human-machine cooperation (Schaeffer et al., 2016). Thus, a likely reason for the relationship is due to the trust aspect of agreeableness.

Agreeableness is also linked with human-human teaming. Among the Big Five personality traits, agreeableness has the strongest correlation with teamwork across professions and job types (Barrick, Mount, & Judge, 2001). In human-human teams, intra-team trust has been linked to team performance, especially for tasks with higher interdependence between teammates (De Jong, Dirks, & Gillespi, 2016).

Trust development will likely be different in human-autonomous agent teams, as the human might not always understand the information and decision-making processes used by the agent, less rich communication media, and different affective responses. Developing the trust required for effective team performance will be especially difficult in these circumstances. Similar to what has been found for virtual teams (Breuer, Hüffmeier, & Hertel, 2016), a high level of dispositional trust is likely to be required for successful performance when working with highly reliable automated agents. Indeed, there is some evidence that people with a propensity to trust humans are also likely to trust automated agents during initial stages of performance (Huang & Bashir, 2017). Thus, agreeableness is likely important to performance when working with autonomous systems.

Openness to Experience. We were not surprised to find near-zero relationships for openness to experience. There are not compelling theoretical links to performance with automation. Facets related to indecisiveness (described earlier) might have stronger relationships with multitasking than is reflected by global openness.

6.7.3 Other Characteristics.

BP and VGE both had strong relationships with our outcomes of interest.

BP. The strong correlation for BP is likely due to its association with vigilance and attention, which can be difficult to maintain when manual task completion is replaced or supplemented with automation. This results in long periods of monitoring and low task-load, but maintaining awareness is still critical to detect automation errors or adverse events.

VGE. The results for VGE were not surprising based on the attention this area has received. There is ample evidence linking hours of video game play to cognitive skills such as working memory (Green & Bavelier, 2003). There is the belief that VGE improves cognitive skills, however there are methodological weaknesses of these studies, and inaccurate descriptions of these studies in the subsequent literature. Another explanation with some support in the literature is that VGE is negatively related to stress during supervisory control tasks (Lin et al., 2017). There are likely multiple underlying constructs driving this relationship. Understanding the underlying attributes driving VGE's prediction will allow us to identify the actual constructs reflected by this indicator.

Opportunities to improve this area of understanding have been recently identified (e.g., Sala et al., 2018; Simons et al., 2016; Waris et al., 2019). Unsworth et al. (2015) argued that the extreme groups and practice effects used in many experimental designs have resulted in overestimated relationships between VGE and cognitive ability. This conclusion not been without controversy. Thus, there is a clear need to explore measurement methods in this area. It should be noted that there are also no validated or widely-used measure of VGE, or an agreed-upon definition of action video games.

PAS. The results for PAS show the potential of this construct to predict important outcomes, especially because of its high relevance to the performance domain. Our findings provide support to the two-dimensional structure of the construct. The dimensions are moderately correlated, yet have different relationships with theoretically-relevant outcomes. This could of course be due to differences in the quality of items for these dimensions rather than due to the true nature of the construct. That is, the all or none dimension items might not have strong correlations due to low reliability or lack of variance.

The most pressing issue in this area is the continued development of PAS measures. The limited studies in this area either only administer items related to one dimension of PAS, or remove certain items. This makes it difficult to compare results across studies. Studies should also report the psychometrics of the PAS measure, since this can aid the interpretation of their findings (i.e., whether low correlations reflect a true relationship or are attenuated due to low reliability or low variance). Item Response Theory could also be a useful tool in evaluating PAS measures and improving variance of the measure (i.e., by creating low and medium difficulty items).

6.8 Future Research

6.8.1 Adopt Reliable Personality Measures.

A major problem discovered during the literature review process was the widespread use of abbreviated personality measures throughout the automation literature (e.g., the Ten Item Personality Inventory; Gosling, Rentfrow, & Swann, 2003). This problem is exacerbated by the widespread failure to analyze and/or report reliabilities (e.g., Cronbach's alpha) in these studies.

While there are conceptual arguments for the benefits of short measures that eschew the conventional wisdom (e.g., Fisher, Mathews, & Gibbons, 2016), the use of unreliable personality measures is almost certainly attenuating the relationships found. This problem is exacerbated by the failure of many authors to report reliabilities, which makes it difficult to estimate or correct for this attenuation.

Indeed, it has been shown that the role of personality in performance is substantially underestimated by short measures (Credé, Harms, Niehorster, & Gaye-Valentine 2012). Thus, these measures are not well-suited to identifying the role of personality in performance. As a consequence, much of the extant literature's findings regarding individual differences and performance with automated systems should not be taken as strongly indicative of the true nature of these relationships.

Hopefully researchers more carefully weigh the advantages of using validated measures against the small increase in time requirements per participant. This will contribute useful knowledge to the literature, even for studies in which individual differences are not the main focus.

6.8.2 Explore Operator States' Relationships with Performance.

Our meta-analysis and literature review identified linkages between individual differences and operator states relevant to performance when working with autonomous systems. Building greater understanding of operator states during performance will improve the external validity of findings to future systems and different job types. An important step in this is to improve self-report measures of operator states. Other methods of measuring operator states (e.g., physiological measures) also should be explored.

6.8.3 Explore the Utility of Multitasking and Monitoring Tests for Selection.

Although theoretical linkages are strong, little research has directly tied tests of multitasking and monitoring/vigilance to performance when working with autonomous systems. Such tests are available commercially or might already exist in DoD test batteries. They have potentially strong criterion-related validities due to their overlap with the performance domain, and their relationships with important individual differences and operator states (e.g., working memory, stress).

6.8.4 Explore Potential Moderators.

The relationships between individual differences and performance are likely moderated by features of the automation (e.g., reliability, anthropomorphism), task environment (e.g., type of simulator, lab versus field study, and type of tasks performed by human), performance criteria (e.g., accuracy, response time). For instance, the criterion-related validity of dispositional trust is likely to depend on the reliability of the automation. BP is likely more predictive in contexts with long periods of monitoring with little action. Testing for moderators will require larger samples, and/or a sufficient number of studies to meta-analytically test for moderation.

7.0 REFERENCES

Amthauer, R., Brocke, B., Liepman, D., & Beauducel, A. (2001). *Intelligence-Structure-Test 2000 R-Handbook*. Hogrefe, Goettingen, Germany.

Atkinson, D., Hancock, P., Hoffman, R. R., Lee, J. D., Rovira, E., Stokes, C., & Wagner, A. R. (2012). Trust in Computers and Robots: The Uses and Boundaries of the Analogy to Interpersonal Trust. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 303–307.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Bagheri, N., & Jamieson, G. A. (2004, October). The impact of context-related reliability on automation failure detection and scanning behaviour. In *IEEE International Conference on Systems, Man and Cybernetics, 1*, 212-217.

Bahner, J. E., Hüper, A. D., & Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies*, *66*, 688-699.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection & Assessment*, *9*(1), 9-30.

Barron, L. G., & Rose, M. R. (2017). Multitasking as a predictor of pilot performance: Validity beyond serial single-task assessments. *Military Psychology, 29,* 316-326. DOI: https://doi.org/10.1037/mil0000168

Baumeister RF, Vohs KD. (2007). *Encyclopedia of Social Psychology*. Thousand Oaks, CA: Sage

Becker, A., Mandell, A.R., Tangney, J.P., Chrosniak, L.D., & Shaw, T.H. (2015). The effects of self-control on cognitive resource allocation during sustained attention: A transcranial Doppler investigation. *Experimental Brain Research*, *233*, 2215–2223. doi:10.1007/s00221-015-4291-z

Behm-Morawitz, E. (2017). Examining the intersection of race and gender in video game advertising. *Journal of Marketing Communications*, *23*(3), 220–239.

Bhana, H. (2010). Trust but verify. *AeroSafety World*, *5*, 13-14.

Billings, C. E., Lauber, J. K., Funkhouser, H., Lyman, G., & Huff, E. M. (1976). *Aviation Safety Reporting System* (Technical Report TM-X-3445). Moffett Field, CA: National Aeronautics and Space Administration Ames Research Center.

Bratfisch, O., Hagman, E., & Puhr, U. (2002). *Simultaneous capacity (Release 23.00): Manual*. Mödling, Austria: Schuhfried.

Breuer, C., Hüffmeier, J., & Hertel, G. (2016). Does trust matter more in virtual teams? A Meta-analysis of trust and team effectiveness considering virtuality and documentation as moderators. *Journal of Applied Psychology*, *101* (8), 1151-1177.

Bühner, M., König, C. J., Prick, M., & Krumm, S. (2006). Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance. *Human*

Performance, 19, 253–275. https://doi.org.ezproxy.libraries.wright.edu/10.1207/s15327043hup1903_4

Burke, J. L., & Murphy, R. R. (2004). Human-robot interaction in USAR technical search: Two heads are better than one. In *Proceedings of the 13th International Workshop on Robot and Human Interactive Communication* (RO-MAN), 2004.

Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., & Kendall, D. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology, 91*(6), 1189-1207.

Button, K. (2009). Multi-aircraft control of UAVs by a single pilot won't come easily, experts warn. *C4ISR Journal*, *8*, 30-32.

Carretta, T.R., & Ree, M. J. (2003). Pilot selection methods. In P. S. Tsang & M. A. Vidulich (Eds.), *Human factors in transportation. Principles and practice of aviation psychology* (pp. 357-396). Mahwah, NJ: Erlbaum.

Casner, S.M., & Schooler, J.W. (2015). Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Consciousness and Cognition*, 35, 33–41.

Chancey, E. T., Bliss, J. P., Yamani, Y., & Handley, H.A.H. (2016). Trust and the compliance–reliance paradigm: The effects of risk, error bias, and reliability on trust and dependence. *Human Factors, 59*(3), 333–345. https://doi.org/10.1177/0018720816682648

Chen, G., Whiteman, J.-A., Gully, S. M., & Kilcullen, R. N. (2000). Examination of Relationships among trait-like individual differences, state-like individual differences, and learning performance. *Journal of Applied Psychology*, *85*(6), 835–847.

Chen, J.Y.C. (2018). Human-autonomy teaming in military settings. *Theoretical Issues in Ergonomics Science*, *19*(3), 255-258.

Chen, J.Y.C., Barnes, M. J., Harper-Sciarini, M. (2011). Supervisory control of multiple robots: Human performance issues and user interface design. *IEEE Transactions on Systems, Man, and Cybernetics–Part C: Applications and Reviews, 41*, 435–454.

Chen, J.Y.C., & Barnes, M. J. (2012). Supervisory control of multiple robots: Effects of imperfect automation and individual differences. *Human Factors*, *54*, 157–174. doi:10.1177/0018720811435843

Chen, J.Y.C., Lakhmani, S. G., Stowers, K., Selkowitz, A. R., Wright, J. L. & Barnes, M. J. (2018). Situation awareness-based agent transparency and human-autonomy teaming effectiveness. *Theoretical Issues in Ergonomic Science, 19 (3),* 259–282, doi: 10.1080/1463922X.2017.1315750.

Chen, J.Y.C., & Terrence, P. I. (2009). Effects of imperfect automation on concurrent performance of military and robotics tasks in a simulated multi-tasking environment. *Ergonomics, 52*, 907–920.

Chiou, E. K., & Lee, J. D. (2016). Cooperation in human-agent systems to support resilience: A microworld experiment. *Human Factors*, *58*(*6*), 846–863.

Clarke, R. I., Lee, J. H., & Clark, N. (2017). Why video game genres fail: A classificatory analysis. *Games and Culture*, *12*(*5*), 445–465.

Claypoole, V.L., Neigel, A.R., Waldfogle, G.E., & Szalma, J.L. (2019). Evaluative social presence can improve vigilance performance, but vigilance is still hard work and is stressful. *JEP: Human Perception and Performance*.

Claypoole, V. L., & Szalma, J. L. (2018). Independent coactors may improve performance and lower workload: Viewing vigilance under social facilitation. *Human Factors*, *60*(6), 822-832.

Colom, R., Martínez-Molina, A., Shih, P., & Santacreu, J. (2010). Intelligence, working memory, and multitasking performance. *Intelligence, 38,* 543–551.

Correia, F., Mascarenhas, S., Prada, R., Melo, F.S., & Paiva, A. (2018). Group-based emotions in teams of humans and robots. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 261-269.

Cox-Fuenzalida, L. E. (2000). *Effect of workload history on vigilance performance*. Unpublished doctoral dissertation, University of Oklahoma, Norman.

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology*, *102*(4), 874-888.

Cumming, R. W., & Croft, P. G. (1973). Human information processing under varying task demand. *Ergonomics*, *16*, 581–586.

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.

Dale, G., & Green, C.S. (2017). The changing face of video games and video gamers: Future directions in the scientific study of video game play and cognitive performance. *Journal of Cognitive Enhancement*, *1* (*3*), 280-294.

Dawes, R. M. (1994). *House of Cards: Psychology and Psychotherapy Built on Myth*, Free Press, NY.

De Jong, B. A., Dirks, K. T., & Gillespie, N. (2016). Trust and team performance: A meta-analysis of main effects, moderators, and covariates. *The Journal of Applied Psychology*, *101*(8), 1134–1150. https://doi-org.ezproxy.libraries.wright.edu/10.1037/apl0000110

de Visser, E. J., Monfort, S. S., Goodyear, K., Lu, L., O'Hara, M., Lee, M. R., ... & Krueger, F. (2017). A little anthropomorphism goes a long way: Effects of oxytocin on trust, compliance, and team performance with automated agents. *Human Factors*, *59*, 116-133.

de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost Human: Anthropomorphism Increases Trust

Resilience in Cognitive Agents. *Journal of Experimental Psychology: Applied, 22*(3), 331–349. https://doi-org.ezproxy.libraries.wright.edu/10.1037/xap0000092

de Visser, E., Shaw, T., Mohamed-Ameen, A., & Parasuraman, R. (2010). Modeling human-automation team performance in networked systems: Individual differences in working memory count. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 54*(14), 1087–1091. https://doi.org/10.1177/154193121005401408

Defense Science Board. (2012). *Task Force Report: The Role of Autonomy in DoD Systems*. Washington, DC: Office of the Undersecretary of Defense for Acquisition, Technology, and Logistics.

Degani, A. (2004). *Taming HAL: Designing interfaces beyond 2001*. NY: Springer.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111*(2), 225–236. https://doi.org/10.1037/0021-843X.111.2.225

Dillard, M. B., Warm, J. S., Funke, G. J., Nelson, W. T., Finomore, V. S., McClernon, C. K., Eggemeier, F.T., Tripp, L.D., & Funke, M. E. (2019). Vigilance tasks: Unpleasant, mentally demanding, and stressful even when time flies. *Human Factors*, *61*, 225-242. https://doi.org/10.1177/0018720818796015

Duckworth, A., Peterson, C., Matthews, M., & Kelly, D. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*. *92*. 1087-101. 10.1037/0022-3514.92.6.1087.

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

Ekstrom, R., French, J., & Harman, H. (1976). *Kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Endler, N., & Parker, J. (1990) Multidimensional assessment of coping: A critical review.

Journal of Personality and Social Psychology, *58*, 844–54.

Endsley, M. R. (1990). Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society Annual Meeting*, *34*(1), 41–45.

Engel, R. W. (2002). Working memory as executive attention. *Current Directions in Psychological Science, 11,* 19-23.

Eschen, C.S.S., Gayraud, K., & Keye-Ehing, D. (2016). Critical Personality Aspects for Human-Machine Interaction in highly automated Aviation. *Presented at the 6th SESAR Innovation Days*, Delft, the Netherlands.

Fahey, K.M., & Miller, M.J. (2017). Unmanned systems integrated roadmap, 2017-2042. *Department of Defense,* Technical Report.

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness: The development and correlates of a new scale. *Journal of Personality Assessment*, *50*, 4-17.

Ferris, T., Sarter, N., & Wickens, C. D. (2010). Cockpit automation: still struggling to catch up, In E. Salas & D. Maurino (2nd Ed.), *Human Factors in Aviation* (pp. 479-503). Burlington, MA: Academic Press.

Fincannon, T., Keebler, J.R., Jentsch, F., & Curtis, M. (2013). The influence of camouflage, obstruction, familiarity and spatial ability on target identification from an unmanned ground vehicle. *Ergonomics*, *56*, 739-751.

Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations*, *46*, 395–417.

Foroughi, C. K., Rovira, E., Rose, K., Davis, D., Jurado, J. J., Chae, J., … Coyne, J. T. (2019). How do individual differences in attention control and working memory capacity relate to monitoring automation? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 1409–1413. https://doi.org/10.1177/1071181319631196

Gambetta, D. (1988). Can we trust trust?, in D. Gambetta (Ed.), *Trust: Making and Breaking Cooperative Relationships* (pp. 213-237. Cambridge: Basil Blackwell.

Geiselman, E. E., Johnson, C. M., Buck, D. R., & Patrick, T. (2013). Flight deck automation: A call for context-aware logic to improve safety. *Ergonomics in Design*, *21*, 13-18.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, F. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, *7*, 7-28. Tilburg, The Netherlands: Tilburg University Press.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40,* 84-96.

Goldberg, D. R., & Stewart, M. R. (1980). Memory overload or expectancy effect? ‘Hysteresis’ revisited. *Ergonomics*, *23,* 1173–1178.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-528.

Gouraud J., Delorme A., & Berberian B. (2018). Influence of automation on mind wandering frequency in sustained attention. *Consciousness and Cognition*, *66*, 54-64.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534.

Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1465–1478.

Green, C. S., Kattner, F., Eichenbaum, A., Bediou, B., Adams, D. M., Mayer, R. E., & Bavelier, D. (2017). Playing some video games but not others is related to cognitive abilities: A critique of Unsworth et al. (2015). *Psychological Science*, 28.

Gao, F., Cummings, M.L., & Solovey, E.T. (2014). Modeling Teamwork in Supervisory Control of Multiple Robots. *IEEE Transactions on Human–Machine Systems*, *44* (*4*), 441-453.10.1109/THMS.2014.2312391

Gugerty, L., & Brooks, J. (2004). Reference-frame misalignment and cardinal direction judgments: Group differences and strategies. *Journal of Experimental Psychology: Applied*, *10*, 75–88.

Gutzwiller, R.S., Lange, D.S., Reeder, J., Morris, R.L., & Rodas, O. (2015). Human-computer collaboration in adaptive supervisory control and function allocation of autonomous system teams. In: Shumaker, R., Lackey, S. (eds.) *Virtual, Augmented and Mixed Reality. VAMR 2015. Lecture Notes in Computer Science, 9179*, pp. 447–456. Heidelberg: Springer.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J.Y.C., de Visser, E., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human–robot interaction. *Human Factors*, *53*, 517–527.

Held, J. D., & Carretta, T. R. (2013). Evaluation of tests of processing speed, spatial ability, and working memory for use in military occupational classification. *NPRST in Work Report*. Millington, TN: Navy Personnel, Research, Studies and Technology.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, 212, 429–437.

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, 224, 429–436.

Helton, W. S., Shaw, T. H., Warm, J. S., Matthews, G., & Hancock, P. A. (2008). Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety, Stress, & Coping: An International Journal*, *21*, 173–184.

Huang, H.Y., & Bashir, M. (2017). Personal influences on dynamic trust formation in human-agent interaction. In *Proceedings of the 5th International Conference on Human Agent Interaction*, (pp. 233–243). ACM, October 2017

Hunter, A. & Eastwood, J. D. (2018). Does state boredom cause failures of attention? Examining the relations between trait boredom, state boredom, and sustained attention. *Experimental Brain Research*, *236*, 2483-2492.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.

Hunter, J.E., & Hunter, R.F. (1984). *Validity and utility of alternative predictors across studies*. *Psychological Bulletin*, *96*, 72-98.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Kaber, D. B., & Endsley, M. R. (1997). The combined effect of level of automation and adaptive automation on human performance with complex, dynamic control systems. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 205–209). Santa Monica, CA: Human Factors and Ergonomics Society.

Kanfer, R. (1990). Motivation theory and industrial and organizational psychology. In M. D. Dunnette & L. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 75–170). Palo Alto, CA: Consulting Psychologists Press.

Karpinsky, N.D., Chancey, E.T., Palmer, D.B., & Yamani, Y. (2018). Automation trust and attention allocation in multitasking workspace. *Applied Ergonomics, 70*, 194–201.

König, C. J., Bühner, M., & Mürling, G. (2005). Working memory, fluid Intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human Performance*, *18*(*3*), 243-266. doi:10.1207/s15327043hup1803_

Kyllonen, P, C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity*?! Intelligence, 14,* 389-433.

Lasota, P. A., & Shah, J. A. (2015). Analyzing the effects of human-aware motion planning on close-proximity human-robot collaboration. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57* (1), 21–33. doi:10.1177/0018720814565188.

Lee, J. D., & See, K. A. (2004). Trust in automation: designing for appropriate reliance. *Human Factors, 46*(1), 50–80.

Levulis, S. J., DeLucia, P. R., & Kim, S. Y. (2018). Effects of touch, voice, and multimodal input, and task load on multiple-UAV monitoring performance during simulated manned-unmanned teaming in a military helicopter. *Human Factors, 60*(8), 1117–1129. https://doi.org/10.1177/0018720818788995

Lin, J. (2017). *The Impact of automation and stress on human performance in UAV operation*. Unpublished doctoral dissertation, University of Central Florida, Orlando.

Lin, J., Wohleber, R. W., Matthews, G., Chiu, Y.-C. P., Calhoun, G. L., Ruff, H. A., & Funke, G. J. (2015). Video game experience and gender as predictors of performance and stress during supervisory control of multiple unmanned aerial vehicles. In *Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting* (pp. 746–750). Santa Monica, CA: Human Factors and Ergonomics Society.

Loft, S., Morrell, D. B., & Huf, S. (2013). Using the situation present assessment method to measure situation awareness in simulated submarine track management. *International Journal of Human Factors and Ergonomics, 2*, 33–48.

Loft, S., Jooste, L., Li, Y. R., Ballard, T., Huf, S., Lipp, O. V., & Visser, T. A. W. (2018). Using situation awareness and workload to predict performance in submarine track management: A multilevel approach. *Human Factors*, *60*(7), 978–991.

Lyons, J.B., & Guznov, S.Y. (2019) Individual differences in human–machine trust: A multi-study look at the perfect automation schema, *Theoretical Issues in Ergonomics Science, 20*(4), 440-458, doi: 10.1080/1463922X.2018.1491071

Lyons, J.B., Ho, N.T., Fergueson, E., Sadler, G., Cals, S., Richardson, C., & Wilkins, M. (2016). Trust of an Automatic Ground Collision Avoidance Technology: A Fighter Pilot Perspective. *Military Psychology*, 28(4), 271-277.

Lyons, J.B., Wynne, K.T., Mahoney, S., Roebke, M.A. (2019). Trust and human-machine teaming: a qualitative study. In W. Lawless, R., Mittu, D., Sofge, L., Moskowitz, & S., Russell, (Eds.), *Artificial Intelligence for the Internet of Everything* (pp. 101-116). Cambridge, MA: Academic Press.

Mandell, A. R., Becker, A., VanAndel, A., Nelson, A., & Shaw, T. H. (2015). Neuroticism and vigilance revisited: A transcranial Doppler investigation. *Consciousness and Cognition*, *36*, 19–26.

Matthews, M. L. (1986). The influence of visual workload history on visual performance. *Human Factors*, *28*, 623–632.

Matthews, G., & Campbell, S.E. (1998). Task-induced stress and individual differences in coping. In *Proceedings of the Human Factors and Ergonomics Society*, *42*, 821-825.

Matthews, G., & Campbell, S.E. (2009). Sustained performance under overload: Personality and individual differences in stress and coping. *Theoretical Issues in Ergonomics Science*, *10*, 417-442.

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

Matthews, G., Warm, J. S., Shaw, T. H., & Finomore, V. S. (2014). Predicting battlefield vigilance: A multivariate approach to assessment of attentional resources. *Ergonomics*, *57*, 856–875.

Matthews, G., Warm, J.S., & Washburn, D. (2007). *Diagnostic methods for predicting performance impairment associated with combat stress*. Report for U.S. Army Medical Research and Materiel Command, contract # W81XWH-04-C-0002.

McKendrick, R., Ayaz, H., Olmstead, R., & Parasuraman, R. (2014). Enhancing dual-task performance with verbal and spatial working memory training: continuous monitoring of cerebral hemodynamics with NIRS. *NeuroImage*, *85*, 1016-1028

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.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196–204. https://doi-org.ezproxy.libraries.wright.edu/10.1037/a0014104

Merritt, S. M. (2011). Affective processes in human–automation interactions. *Human Factors*, *53*(4), 356-370.

Merritt, S., Heimbaugh, H., LaChapell, J., & Lee, D. (2013). I trust it, but I don't know why: Effects of implicit attitudes toward automation on trust in an automated system. *Human Factors*. 55. 520-34. 10.1177/0018720812465081.

Merritt, S. M., Ako-Brew A., Bryant W. J., Staley A., McKenna M., Leone A. & Shirase L. (2019). Automation-induced complacency potential: Development and validation of a new scale. *Front. Psychol.* 10:225. doi: 10.3389/fpsyg.2019.00225

Metzger, U., & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors*, *47*, 35-49.

Mojzisch, A., Krumm, S., & Schultze, T. (2014). Do high working memory groups perform better? A conceptual approach linking individual differences in working memory capacity to group performance. *Journal of Personnel Psychology*, *13*(3), 134–145.

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: vigilance and task complexity effects. *Human Factors*, *38*(2), 311–322.

Moray, N., & Inagaki, T. (1999). Laboratory studies of trust between humans and machines in automated systems. *Transactions of the Institute of Measurement and Control*, 21, 203-211.

Morgan, B., D'Mello, S., Fike, K., Abbott, R., Haass, M., Tamplin, A., Radvansky, G., & Forsythe, C. (2011). Individual differences in multitasking ability and adaptability. In *Proceedings of the Human Factors & Ergonomics Society 55th Annual Meeting* (pp. 919–923). Santa Monica, CA: Human Factors and Ergonomics Society.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167–193.

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, MQ-9 Reaper Operators* (Report no. AFRL-SA-WP-TR-2011-0003). U. S. Air Force School of Aerospace Medicine: Wright-Patterson AFB, OH.

Panganiban, A. R., & Matthews, G. (2014). Executive functioning protects against stress in UAV Simulation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, No.1, pp. 994-998). Sage.

Panganiban, A. R., Matthews, G., & Long, M. D. (2019). Transparency in Autonomous Teammates: Intention to Support as Teaming Information. *Journal of Cognitive Engineering and Decision Making*. https://doi.org/10.1177/1555343419881563

Parasuraman, R., Cosenzo, K. A., & de Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, *21*, 270-297.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381–410. https://doi-org.ezproxy.libraries.wright.edu/10.1177/0018720810376055

Parasuraman, R., Molloy, R. & Singh, I. L. (1993). Performance consequences of automation-induced ‘‘complacency’’. *International Journal of Aviation Psychology*, 3, 1 – 23.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*, 230-253.

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*, 286-297.

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.

Poposki, E. M., Oswald, F. L., & Chen, H. T. (2009). *Neuroticism negatively affects multitasking performance through state anxiety* (NPRST-TN-09-3). Technical report for Navy Personnel Research, Studies, and Technology. Millington, TN.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual: Section 1. General overview*. Oxford, England: Oxford Psychologists Press.

Reinerman-Jones, L. E., Matthews, G., Langheim, L. K., & Warm, J. S. (2011). Selection for vigilance assignments: A review and proposed new direction. *Theoretical Issues in Ergonomics Science, 12*, 273-296.

Richards, D., Izzetoglu, K., & Shelton-Rayner, G. (2017). UAV Operator mental workload - A neurophysiological comparison of mental workload and vigilance. AIAA Modeling and Simulation Technologies Conference 10.2514/6.2017-3670.

Robison, M. K., Miller, A. L., Unsworth, N. (2020). A multi-faceted approach to understanding individual differences in mind-wandering. *Cognition*, 198.

Rose, C. L., Murphy, L. B., Byard, L., & Nikzad, K. (2002). The role of the Big Five personality factors in vigilance performance and workload. *European Journal of Personality*, *16*(3), 185–200. https://doi-org.ezproxy.libraries.wright.edu/10.1002/per.451

Rose, M. R., Barron, L. G. , Carretta, T. R., Arnold, R. D., & Howse, W. R. (2014). Early identification of unmanned aircraft pilots using measures of personality and aptitude, *International Journal of Aviation Psychology*, *24*(1), 36-52.

Sackett, P. R., & Lievens, F. (2008). Personnel Selection. *Annual Review of Psychology*, *59*, 419-450.

Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin*, *144*(*2*), 111-139. http://dx.doi.org/10.1037/bul0000139

Salgado, J., Anderson, N. R., Moscoso, S,. Bertua, C., de Fruyt, F., & Rolland J. P. (2004). A meta-analytic study of GMA validity for different occupations in the European community. *Journal of Applied Psychology, 88*, 1068-1081.

Sato, T., Yamani, Y., Liechty, M., & Chancey, M. (2019). Automation trust increases under high-workload multitasking scenarios involving risk. *Cognition, Technology & Work*, 1-9.

Saqer, H. & Parasuraman, R. (2014). Individual performance markers and working memory predict supervisory control proficiency and effective use of adaptive automation. *International Journal of Human Factors and Ergonomics*, *3* (*1*), 15–31.

Schaefer, K. E., Chen, J.Y.C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, *58*(3), 377–400. https://doi-org.ezproxy.libraries.wright.edu/10.1177/0018720816634228

Schmidt, F.L., & Hunter, J.E. (1998). *The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings*. *Psychological Bulletin, 124*, 298-307.

Schmitt, N. (2014). Personality and cognitive ability as predictors of effective performance at work. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 45-65.

Schmitt, N., & Chan, D. (1998). *Personnel selection: A theoretical approach.* Thousand Oaks, CA: Sage.

Schweizer, K., & Moosbrugger, H. (2004). Attention and working memory as predictors of intelligence. *Intelligence, 32*, 329–347.

Sellner, B., Heger, F.W., Hiatt, L. M., Simmons, R., & Singh, S. (2006). Coordinated multi-agent teams and sliding autonomy for large-scale assembly. *Proceedings of the IEEE*, *94*(7), 1425–1444. Retrieved from https://doiorg.ezproxy.libraries.wright.edu/10.1109/JPROC.2006.876966

Shackman, A. J., Sarinopoulos, I., Maxwell, J. S., Pizzagalli, D. A., Lavric, A., & Davidson, R. J. (2006). Anxiety selectively disrupts visuospatial working memory. *Emotion, 6*(1), 40–61. https://doi-org.ezproxy.libraries.wright.edu/10.1037/1528-3542.6.1.40

Shaw, T.H., Matthews, G., Warm, J.S., Finomore, V., Silverman, L., & Costa, P.T., Jr.

(2010). Individual differences in vigilance: Personality, ability and states of stress.

Journal of Research in Personality, 44, 297-308.

Sheridan, T.B. (2002). *Supervisory control*. Santa Monica, CA: Wiley Interscience.

Sibley, C., Coyne, J. T., & Sherwood, S. (2016). Research considerations and tools for evaluating human automation interaction with future unmanned systems. In W. F. Lawless, R. Mittu, D. Sofge, & S. Russell (Eds.), Autonomy and artificial intelligence: A threat or savior? (pp. 157–178). Cham, Switzerland: Springer.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do “brain-training” programs work? *Psychological Science in the Public Interest, 17,* 103–

186. http://dx.doi.org/10.1177/1529100616661983

Spielberger, C. D. (1985). Anxiety, cognition and affect: A state-trait perspective. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders.* (pp. 171–182). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Retrieved from https://search-ebscohostcom.ezproxy.libraries.wright.edu/login.aspx?direct=true&db=psyh&AN=1985-97708-008&site=ehost-live&scope=site

Stauffer, J. M., Ree, M. J., & Carretta, T. R. (1996). Cognitive-components tests are not much more than *g*: An extension of Kyllonen's analyses. *The Journal of General Psychology, 123*, 193-205.

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*, 71–96. doi:10.1037/a0024170

Tasa, K., Taggar, S., & Seijts, G. H. (2007). The development of collective efficacy in teams: A multilevel and longitudinal perspective. *Journal of Applied Psychology, 92*, 17–27. doi: 10.1037/0021-9010.92.1.17

Thompson, W. T., Lopez, N., Hickey, P., DaLuz, C., & Caldwell, J. L. (2006). Effects of shift work and sustained operations: Operator performance in remotely piloted aircraft (OP-REPAIR). *Presented at the 311th Performance Enhancement Directorate and Performance Enhancement Research Division*, Brooks City-Base, TX.

Thropp, J. E., Oron-Gilad, T., Szalma, J. L., & Hancock, P. A. (2018). Calibrating adaptable automation to individuals. *IEEE Transactions on Human-Machine Systems*, *48*(6), 691–701. https://doi-org.ezproxy.libraries.wright.edu/10.1109/THMS.2018.2844124

United States Air Force (2015). *Autonomous horizons: System autonomy in the Air Force- A path to the future, volume 1: Human-autonomy teaming*, AF-ST-TR-15-01. United States Air Force, Office of the Chief Scientist.

Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is Playing Video Games Related to Cognitive Abilities? *Psychological Science*, *26*(6), 759–774.

Unsworth, N., & Robison, M. K. (2020). Working memory capacity and sustained attention: A cognitive-energetic perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(*1*), 77–103.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2009). Exploring the relations among working memory capacity, attention control, and fluid intelligence from a dual-component framework. *Psychology Science Quarterly, 51*, 388–402.

Wang, N., Pynadath, D.V., & Hill, S.G. (2016). Trust calibration within a human-robot team: Comparing automatically generated explanations. *Presented at the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Christchurch.

Waris, A., Mendez, I., Englehart, K., Jensen, W., & Kamavuako, E. N. (2019). On the robustness of real-time myoelectric control investigations: a multiday Fitts' law approach. *Journal of Neural Engineering*, *16*(2), 026003. https://doi-org.ezproxy.libraries.wright.edu/10.1088/1741-2552/aae9d4

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: theory and applications* (pp. 183–200). Mahwah, NJ: Erlbaum.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*, 433-441.

Waters, S., Russell, T., Shaw, M., Allen, M., Sellman, W., & Geimer, J. (2009). *Development of a methodology for linking ASVAB content to military job information and training curricula*. (HumRRO Final Report FR-09-64). Alexandria, VA: Human Resources Research Organization.

Wickens, C. D., Clegg, B. A., Vieane, A. Z., & Sebok, A. L. (2015). Complacency and Automation Bias in the Use of Imperfect Automation. *Human Factors*, *57*(5), 728–739.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2016). *Engineering Psychology & Human Performance* (4^{th} Ed.). NY: Taylor & Francis.

Wickens, C. D., Li, H., Santamaria, A., Sebok, A., & Sarter, N. B. (2010). Stages and levels of automation: An integrated meta-analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, No. 4, pp. 389-393). Sage CA: Los Angeles, CA: Sage.

Wiener, E. L. (1981). Complacency: Is the term useful for air safety? In *Proceedings of the 26th Corporate Aviation Safety Seminar* (pp. 116–125). Denver, CO: Flight Safety Foundation

Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: Promises and problems. *Ergonomics*, *23*, 995-1011.

Wohleber, R. W., Matthews, G., Lin, J., Szalma, J. L., Calhoun, G. L., Funke, G. J., Chiu, P., Ruff, H. A. (2019). Vigilance and automation dependence in operation of multiple unmanned aerial systems (UAS): A simulation study. *Human Factors, 61*(3), 488–505.

Wright J.L., Chen J.Y., Barnes M.J. (2018). Human–automation interaction for multiple robot control: the effect of varying automation assistance and individual differences on operator performance. *Ergonomics*, *61*(8):1033–1045.

Wright, J. L., Chen, J. Y., Barnes, M. J., & Hancock, P.A. (2017). *Agent reasoning transparency: The influence of information level on automation induced complacency* (No. ARL-TR-8044). US Army Research Laboratory Aberdeen Proving Ground United States.

Wynne, K. T., & Lyons, J. B. (2018). An integrative model of autonomous agent teammate-likeness. *Theoretical Issues in Ergonomic Science*, *19*(3), 353-374.

Zaccaro, S. J. (2007). Trait-based perspectives of leadership. *American Psychologist*, *62*(1), 6-16. doi:10.1037/0003-066X.62.1.6

8.0 LIST OF SYMBOLS, ACRONYMS AND ABBREVIATIONS

AC	Attentional Control
AoN	All-or-None
BP	Boredom Proneness
CI	Confidence Interval
CITS	Coping Inventory for Task Situations
DoD	Department of Defense
GMA	General Mental Ability
HAI	Human-Automation Interaction
HE	High Expectations
HMT	Human-Machine Teaming
LL	Lower Limit
NASA-TLX	National Aeronautics and Space Administration Task Load Index
PAS	Perfect Automation Schema
RAPM	Raven Advanced Progressive Matrices
RPA	Remotely Piloted Aircraft
SA	Situation Awareness
SAOC	skills, abilities, and other characteristics
SFQ	Cognitive Failures Questionnaire
SpA	Spatial Ability
UAV	Unmanned Aerial Vehicle
UGV	Unmanned Ground Vehicle
UL	Upper Limit

USAF	United States Air Force
VGE	Video Game Experience
WM	Working Memory