January 2021

An Assessment Of Remotely Piloted Aircraft Training Methods And Measures

Zachary Waller

AN ASSESSMENT OF
REMOTELY PILOTED AIRCRAFT
TRAINING METHODS AND MEASURES

by

Zachary Peter Waller
Bachelor of Science, University of North Dakota, 2012
Master of Science, University of North Dakota, 2013

Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, ND
December
2021

This dissertation, submitted by Zachary P. Waller in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

____

Dr. Robert Stupnisky, Chairperson

____

Dr. Virginia Clinton-Lisell

____

Dr. Ryan Summers

____

Dr. Thomas Petros

This dissertation is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

____

Chris Nelson
Dean of the School of Graduate Studies

____

Date

PERMISSION

Title	An Assessment of Remotely Piloted Aircraft Training Methods and Measures
Department	Education Foundations and Research
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in his absence, by the chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Name: Zachary P. Waller

Date: 15 Sept, 2021

Table of Contents

Abstract iv

Introduction to Program of Research 1

I. Waller et al. (2016) 7

II. Waller and Stupnisky (Manuscript) 31

III. Waller and Petros (Manuscript) 51

Discussion & Conclusions 73

References 80

Appendix A Codebooks for Studies I-III A-1

Codebook for Study I A-1

Codebook for Study II A-5

Codebook for Study III A-9

List of Figures

Study I

[NONE]

Study II

Figure 1, Estimated Marginal Mean Performance by Modality 43

Study III

Figure 1, Simulated Flight Pattern 61

Figure 2, Cognitive State Metric for Workload:

ABM's High Workload Metric during Checklist and Flight Pattern Events 64

Figure 3, Cognitive State Metric for Engagement:

ABM's High Engagement Metric during Checklist and Flight Pattern Events 64

List of Tables

Study I

Table 1, Research Design 19

Table 2, Descriptive Results According to Instructional Method......................... 20

Table 3, One Way ANOVA Results According to Instructional Method 21

Table 4, One Way ANOVA results According to Pilot Certification................... 22

Table 5, Descriptive Results of Learner Attitude 22

Study II

Table 1, Research Design 40

Table 2, Descriptive Statistics for Student Performance 41

Table 3, Comparison of Self-Reported Hours of Study...................................... 42

Study III

Table 1, Probability of High Engagement and Workload.................................... 63

Abstract

The collegiate aviation programs of higher education are seeking to adapt their capabilities and expertise toward educating a generation of airmen who will operate Remotely Piloted Aircraft (RPA). The collection of studies presented in this dissertation address this interest as higher education programs investigate the value of modalities and pedagogies, tune the application of instructional aids, and assess novel measurements for how students interact with their training. Three studies were completed in building this program of research. Study I, Waller et al. (2016), was published in a peer-reviewed journal and is adapted for reprint with permission. Study I established the effectiveness of a software trainer to improve students' ability to interact with the MQ-9 Remotely Piloted Aircraft (RPA) when students were granted access via either a traditional, blended, or distance modality. Study II expanded the work of Waller et al. (2016), increasing the sample size to reach across the curriculum as well as accounting for students' pilot certification to better isolate the effect of modality on student performance with the MQ-9 Heads Down Display (HDD) menus. Lastly, Study III assessed whether workload and engagement could be measured by cognitive state estimation as students conducted simulated MQ-1 RPA training. This program of research advances the understanding of RPA instruction by (1) assessing tools and methods that can contribute to a student's training, and (2) demonstrating that cognitive state measurement is sensitive to changes in student workload and engagement.

Keywords: remotely piloted aircraft, blended learning, electroencephalograms, workload, cognitive states

Introduction to Program of Research

The University of North Dakota (UND) and its John D. Odegard School of Aerospace Science offers a variety of educational opportunities in aviation. One of the world's largest civilian training fleets, an estate of facilities, well-vetted policies, and a growing multimedia capability have allowed students to explore majors in commercial aviation, flight education, Air Traffic Management (ATM), aviation management, and one of the nation's first majors in Unmanned Aircraft Systems (UAS) Operations (Miller, 2019).

The opportunities presented by unmanned aircraft have long since captured the attention of aviation educators and researchers in higher education, regulatory administrations, entrepreneurs in industry, as well as state and congressional legislatures (Banks et al., 2018; Jenkins & Vasigh, 2013; Miller, 2019; North Dakota Aeronautics Commission, 2010, 2015). Today, the collegiate aviation programs in higher education leverage their experience across the aeronautical sciences toward supplying the newest generation of airmen to the unmanned aircraft industry (Waller & Bridewell, 2014). Like aspiring airline pilots (Lutte & Lovelace, 2016), students of unmanned aircraft bear high costs for flight training. The program of research below assesses trainers, blended learning, and measures of human performance as specific flight training adaptations. Reducing contact time in a flight training device – which has a high operational cost – or tailoring training to an individual's competency and performance are opportunities which can improve both the quality and efficiency of training.

Overview of Study I

Published in the Journal of Unmanned Aerial Systems in 2016, Waller et al. conducted a pilot study collecting data on the application of a Heads Down Display (HDD) Menu Trainer. While decades of experience and human factors study have informed the design of control interfaces in modern manned aircraft, the flight control interfaces unique to unmanned aircraft have been associated with several mishaps and accidents (Williams, 2004, 2006). The nested HDD menus of the MQ-1 and MQ-9 are used for many functions and the unfamiliar interface was identified as problematic during student training. A software trainer was designed in response to these challenges through the cooperative efforts of the University of North Dakota and the Air Force Research Laboratory (AFRL). The HDD menu trainer was designed to familiarize students with the layout and manipulation of the HDD menus for the MQ-1 and MQ-9 RPA. Waller et al. assessed students' ability to interact with the menus after using the trainer and also investigated whether student performance with the trainer would vary when it was not delivered as part of a traditional lesson with an instructor.

A mixed ANOVA compared pretest and posttest scores ($n = 15$) across modalities (i.e. traditional, blended, and distance pedagogies). Results demonstrated that the trainer significantly improved the performance of all students ($p < 0.001$), however no significant effect was found between the different modalities. Although there were no significant differences noted in pretest or posttest scores between methods of instruction, it was observed that students holding commercial pilot certificates performed significantly higher on the pretest that those with no FAA pilot certification ($p < 0.05$). The study demonstrated the HDD menu trainer's capacity to improve students' navigation and manipulation of the MQ-9 menu structure but recommended controlling for pilot certification as its effects across instructional methods were investigated further.

Overview of Study II

Extending the work of Waller et al. (2016), this study further investigated differences in student performance on the HDD menu trainer when it is applied in traditional, blended, and distance modalities. This study was designed to allow for improved isolation of the variation in performance that could be uniquely attributed to modality. As pilot certification had been shown to affect performance with the trainer, a sample of students (n = 102) both with and without FAA pilot certification completed the same pretest and posttest evaluation used by Waller et al. Within this sample, 26 participants held no FAA pilot certificate, 48 participants held a Private certificate, and 27 participants carried Commercial certification. Students participating in the blended (n = 29) and distance (n = 30) modalities, who accessed the HDD menu trainer remotely, were also asked to self-report the hours spent studying with the trainer on their own.

Level of pilot certification was entered as a covariate in a mixed factorial ANCOVA. Results illustrated – again – the effectiveness of the HDD menu trainer with a main within-subjects effect on performance $F(1,93)=27.65$, $p<.001$. That is, independent of both the modality and pilot certification, posttest scores were higher than pretest scores. Regardless of modality, student performance was – once more – higher for students holding an FAA pilot certificate than for those without $F(1,93)=3.97$, $p<.05$. No significant difference was found between the hours of study reported by participants in blended and distance instruction $t(54)=-0.08$, *ns*, and neither modality nor pilot certification was found to significantly moderate student performance from pretest to posttest. Strictly in terms of performance with the trainer, the blended and distance instructional methods performed at least as well as the traditional instructional method.

The educational process should not be reduced to a single metric or measurement. However, these results on performance illustrate how the HDD menu trainer can be utilized in a way which compliments practical one-on-one flight instruction – a hallmark of aviation education – without compromising on students' ability to navigate and manipulate the HDD menus for the MQ-1 and MQ-9 RPA. Today it has become a standalone staple of upper-division instruction for these aircraft.

Overview of Study III

This study - The Effectiveness of Operator State Monitoring in Measuring Remotely Piloted Aircraft (RPA) Training – explores whether electroencephalogram (EEG) technology is able to measure changes in cognitive workload and engagement in remote pilots during their simulated RPA training. Cognitive workload and task engagement are common constructs in human performance research, and represent the supply-demand relationship of cognitive resources and the attentional resources available to attend a task, respectively (Bernhardt et al., 2019). Over the past two decades, equipment and indices have been developed to measure these constructs of performance in laboratory settings using basic cognitive tasks (Berka et al., 2007). These measures have been proposed for assessing the effectiveness of training and simulation programs because they are able to assess change in cognitive state which is not obvious from task performance alone (Berka et al., 2007; Parasuraman, 2015). Electroencephalograms have been shown to reflect workload levels and sustained attention during training and learning, however a limited number of studies have examined the effectiveness of EEG in operational settings (Bernhardt et al., 2019; Mathan & Yeung, 2015; Mills et al., 2017; Yuan et al., 2014).

Extending the technology's application in other areas of aviation (Aricò et al., 2016; Bernhardt et al., 2019; Borghini et al., 2015), EEG data was collected from remote pilots (n = 10) during simulated MQ-1 RPA training events in the PRINCE device. Posterior probabilities of ABM's high workload and engagement metrics were collected throughout the simulation. Results demonstrated that EEG-based cognitive state metrics are able to detect subtle changes in operator workload during simulated RPA operations. The NASA TLX was administered to collect a subjective measure of workload but no significant association was observed between the subjective and EEG-based measures of workload.

Also noted was significantly reduced workload during those legs of the flight pattern assisted by the heading hold function of the autopilot than for those legs where remote pilots were unassisted by this automation. In addition to proposing these metrics for measuring the effect of training over time, this relationship between remote pilot cognitive workload and autopilot use could also justify design of a procedure which investigates the impact of automation on workload and other cognitive states during simulated RPA operations.

Purpose of the Research Program

Each of the three studies presented investigated methods and measurement of training in unmanned aircraft – recognized more broadly as Remotely Piloted Aircraft (RPA). When a study reflects peer-reviewed work previously published or submitted, the reprint is accompanied by reference to the associated journal with permission to reprint for the purposes of this program. References are detailed within each study but a summary of references across the entire program of research also precedes the appendices. Appendix A contains annotated summaries of the dataset variables for each effort. Results and conclusions are offered in each study, however, a Discussion and Conclusions chapter below offers a synthesized perspective of the research program and also aligns these methods and measures with opportunities for continued study.

Study I
Waller et al. (2016)

**Medium Altitude Long Endurance RPA Training:
A Pilot Study in Blended Learning**

Abstract

Since April of 2011, research and development efforts between the Air Force Research Laboratory (AFRL) and the University of North Dakota (UND) have progressed through the "Science and Technology for Warfighter Training and Aiding." Cooperative Agreement. One product of these cooperative efforts has been a Heads Down Display (HDD) Menu Trainer. Designed to familiarize students with the layout and manipulation of the HDD menus for either the MQ-1 or MQ-9, a parallel pretest/posttest design was designed to examine the efficacy of this HDD menu trainer as training aid in traditional, blended, and distance pedagogies.

Results of a mixed ANOVA indicated the trainer significantly improved performance from pretest to posttest scores across all groups ($p<0.001$), however comparing these scores according to instructional intervention (i.e. Traditional, Blended, and Distance) found no significant effect. No significant differences were observed between pretest, posttest, or percent change scores according to instructional intervention. Analysis of the same variables with respect to pilot certification revealed that learners holding a Commercial pilot certificate scored significantly higher on the pretest than those with no FAA (Federal Aviation Administration) pilot certification ($p<0.05$), and learners with no FAA pilot certificate demonstrated significantly higher percent changes from pretest to posttest than learners with Commercial pilot certificates ($p<0.05$). While, it is clear that the HDD menu trainer has demonstrated effectiveness in improving a student's ability to navigate and manipulate the MQ-9 menu structure, the subtle differences between instructional methods will require further investigation. Future studies are encouraged to investigate the benefits and effectiveness of each instructional method while controlling for pilot certification.

Medium Altitude Long Endurance RPA Training:
A Pilot Study in Blended Learning

In the past two decades, the availability and capability of computer technologies have greatly expanded the educational options available to learners and instructors alike (Osguthorpe & Graham, 2003). Integrating these advances into pedagogy, which recognizes and capitalizes on the inherent strengths of both traditional (i.e. face-to-face) and distance systems of delivery, is the challenge that blended learning offers. Computer Based Training (CBT) modules offer a specific and contemporary example of these expanded educational options, and have been defined as "… self-contained, interactive, often asynchronous, computer-based program[s] designed for self-paced instruction that uses features of learner control coupled with predesigned material, required responses and feedback" (Bedwell & Salas, 2010, p. 240).

Statement of the Problem

Since April of 2011, research and development efforts between the Air Force Research Laboratory (AFRL) and the University of North Dakota (UND) have progressed through the "Science and Technology for Warfighter Training and Aiding." Cooperative Agreement. This CA (FA8650-11-2-6212), is producing a state-of-the-art curriculum for Medium Altitude, Long Endurance Remotely Piloted Aircraft (MALE RPA) pilots and sensor operators, as well as establishing infrastructure for future research efforts. One product of these cooperative efforts has been a Heads Down Display (HDD) Menu Trainer. This CBT module, developed by UND's Aerospace Network, was designed to familiarize students with the layout and manipulation of the HDD menus for either the MQ-1 or MQ-9.

The efficacy of the HDD menu trainer to improve a student's ability in navigating and manipulating the MQ-9 menu structure, as well as its application as training aid in blended pedagogy, or standalone teaching tool in distance pedagogy have not yet been examined. This need for evaluative validation fits well into gaps in extant literature regarding Computer Aided Instruction (CAI) (Adler & Johnson, 2000). In characterizing literature related to CAI, Adler and Johnson (2000) concluded that evaluation articles on the topic remain uncommon in comparison to demonstrations and media-comparative studies, and call for future research to be more aware of these gaps if CAI literature is to mature.

Purpose of the Study

The purpose of this pilot study was to examine the expertise of students in navigating and manipulating the Heads-Down Display (HDD) menus of MALE RPA when provided either traditional, blended, or distance instruction. Learner knowledge gains between groups were measured by both pretest and posttest assessments to assess the effectiveness (1) of the HDD menu trainer, and (2) its potential for use in a variety of instructional methods.

Literature Review

Blended learning

While used frequently throughout academic journals and conferences (Osguthorpe & Graham, 2003), a strict definition of blended learning appears elusive in the extant literature. In his work describing the definitions and directions of blended learning environments, Osguthorpe (2003) offered that,

> "Blended learning combines face-to-face with distance delivery systems… the internet is involved, but it's more than showing a page from a website on the classroom screen. And it all comes back to teaching methodologies – pedagogies that change according to the unique needs of learners. Those who use blended learning environments are trying to maximize the benefits of both face-to-face and online methods – using the web for what

it does best, and using class time for what it does best." (Osguthorpe & Graham, 2003, p. 227)

Osguthorpe and Graham (2003) stress that blended approaches are based upon the assumption that inherent benefits, and weaknesses, exist for both face-to-face interaction and distance delivery. Educators employing blended approaches to instruction must discern the best balance between online access to knowledge and face-to-face human interaction as they develop each course (Osguthorpe & Graham, 2003). Evaluative works on curricula which fall under Osguthorpe and Graham's (2003) working definition of blended learning strategies, or are specific to Bedwell & Salas (2010) definition of CBT, are reviewed in the sections that follow. These efforts have been organized according to domains regarding (1) knowledge gains, (2) learner attitudes, and (3) learning efficiency, as offered by Chumley-Jones, Dobbie, and Alford (2002).

Learner knowledge gains

Efforts addressing learner knowledge gains have assessed change in participant performance as a result of intervention with some manner of computer assisted, or computer based instruction. The majority of studies in this domain measured change using multiple choice test-scores. Pretest/posttest self-controlled studies were the most common design, however others such as self-selected controlled studies, assigned crossover trials, and randomized controlled trials methodologies were also noted (Chumley-Jones et al. 2002). Several within-group methodologies were able to successfully document significant increases in performance as a result of distance instruction (Boyle, Bradley, Chalk, Jones, & Pickard, 2003; Curran, Hoekman, Gulliver, Landells, & Hatcher, 2000; Engel, Crandall, Basch, Zybert, & Wylie-Rosett, 1997; Francis, Mauriello, Phillips, Englebardt, & Grayden, 2000; Harris, Salasche, & Harris, 2001;

Kronz, Silberman, Allsbrook Jr., & Epstein, 2000; Perryer, Walmsley, Barclay, Shaw, & Smith, 2002).

Although within-group assessments of distance instruction were common, between group methodologies allow comparisons to be made across or against alternative pedagogical strategies (i.e. traditional face-to-face, blended, and standalone distance). In these designs, literature which indicated a lack of significant difference in terms of knowledge gains appear to be the majority when distance and traditional pedagogies are compared (Baumlin, Bessette, Lewis, & Richardson, 2000; Bell, Fonarow, Hays, & Mangione, 2000; Block, Felix, Udermann, Reineke, & Murray, 2008; Rivera & Rice, 2002; Rose, Frisby, Hamlin, & Jones, 2000; Sakowski, Rich, & Turner, 2001; Woo & Kimmick, 2000). Allen, Mabry, Mattrey, Bourhis, Titsworth, & Burrell (2004) also found little distinction between traditional and distance learning classrooms on the basis of performance, but offer that no clear decline in educational effectiveness is noted when utilizing distance education technology.

Other between-groups designs did identify significant differences in favor of distance and blended pedagogies. For example, in their examination of potential pedagogic advantages of distance methods of instruction, Lipman, Sade, Glotzbach, Lancaster, and Marshall (2001) compared a traditional classroom course with the same course supplemented by internet-based discussion. Results indicated that performance was higher ($p< 0.005$) in the blended course than the traditional course (Lipman et al. 2001). Melton, Graf, and Chopak-Foss (2009) compared student achievement in blended and traditional pedagogies with mixed results. However, the grades of students in the blended course were found to be significantly higher ($p<0.05$) than those in the traditional course (Melton et al. 2009).

In 2007, Pereira, Pleguezuelos, Meri, Molina-Ros, Molina-Tomas, and Masdeu, examined the efficiency of blended pedagogy, and found that students receiving blended learning received significantly higher grades ($p< 0.0001$) than those in the traditional group (Pereira, et al., 2007). Student feedback also indicated that students felt the course design was an effective (88%) and efficient (92%) method of learning, and helped to familiarize them with resources on the internet (96%) (Pereira, et al., 2007). Further, students' confidence, measured before and after the intervention, showed significant improvement (p <0.001).

Learner attitudes

Since, the late 1990's students have valued the "…flexibility, timeliness, efficiency and breadth of access to relevant information offered by the [internet]" (Agius & Bagnall, 1998, p. 337). Another facet commonly used to evaluate pedagogy, and the second category offered by Chumley-Jones et al. (2002), learner attitudes have been measured and examined regularly in the extant literature.

In their study, Baumlin, et al. (2000) examined course satisfaction with a participant survey. Results indicated that 65% of participants said they wanted computer-assisted instruction as an adjunct to their course curricula, but only 28% of the students with access actually utilized the module. Participants who did use it rated it useful (4.2/5), easy to use (4.4/5), and easy to access (4.1/5). Of the students with access to the online module who chose not to use it, 77.8% reported a lack of time as the reason for not using the module (Baumlin, et al. 2000). In Bell et al. (2000), ratings on a learner satisfaction scale indicated that students using the online tutorial displayed higher satisfaction with the curriculum (Bell et al. 2000).

The 2000 work of Curran et al. also made a general measure of learner attitude. Participants indicated high satisfaction with the self-paced instruction and use of the asynchronous computer conferencing for collaboration among colleagues (Curran et al. 2000). A voluntary satisfaction survey by Harris et al. (2001), indicated extremely high user satisfaction with a distance curriculum. A learner satisfaction survey by Melton et al. (2009), indicated higher satisfaction from students receiving blended learning course delivery ($p < 0.01$). Authors concluded that the blended course delivery was preferred over the traditional lecture format, challenging teachers' traditional approach to delivering general health courses at the university level (Melton et al. 2009). Horsch, Balback, Melnitzki, and Knauth (2000) conducted a simple survey design to measure learner attitudes regarding a distance course. On a scale of 1 to 5, (1=*very good;* 5=*very bad*) students (n = 32) rated the online module at 1.93. In a self-assessment of knowledge gained, 18 of 32 students indicated they had acquired new knowledge, and 10 indicated that learning with the online text was more efficient than learning with a conventional textbook (Horsch et al. 2000).

Hsu and Hsieh (2011) utilized four scales (i.e. the Case Analysis Attitude Scale, Case Analysis Self-Evaluation Scale, Blended Learning Satisfaction Scale, and Metacognition Scale) for students to rate their own performance in blended and traditional delivery courses. Results indicated no difference between groups on any of the self-reported performance scales measured at pretest and posttest. Authors offer that these results demonstrate that both blended learning and traditional classroom lectures are both effective avenues for presenting materials and exchanging ideas to understand course content, and recommend that newly developed course modules and innovative course components should be tested repeatedly for effectiveness (Hsu & Hsieh, 2011). Smyth, Houghton, Cooney, and Casey (2012) interviewed focus groups of students regarding their blended learning experience, and found that students received the blended learning method positively, but offered that the online component meant little time away from study, suggesting that it was more invasive on their everyday life (Smyth et al. 2012).

In their examination of the effectiveness of traditionally and distance courses, Rose et al. (2005) also made a point to measure student satisfaction. No significant differences were reported for (1) communication with classmates, (2) instructor, (3) assignments, (4) review sessions, (5) relevance of course, or (6) the overall course (Rose et al. 2000). Pereira et al. (2007) also observed no statistical difference in overall satisfaction between their blended and traditional courses. Rivera and Rice (2002), who conducted a pilot study evaluating three class formats (i.e. traditional, distance, and blended) found that measures of student satisfaction seemed to indicate that relative to the traditional and blended courses, students in the distance course were less satisfied. Woo and Kimmick (2000) also compared student satisfaction, but found that participants in the distance course reported significantly higher ($p < 0.05$) stimulation of learning compared to those in the traditional lecture course.

As with the efforts addressing learner knowledge gains, measurements of learner attitudes have returned mixed responses. Aside from noting a positive disposition toward pedagogies utilizing some manner of computer assisted, or computer based instruction from the majority of the works, these results are difficult to generalize. While measuring learner attitudes toward experimental curriculums appears commonplace, there seems to be little standardization or congruence in method of measurement.

Learning efficiency

The final and briefest of the three categories examined is learning efficiency. Requiring at minimum a between groups comparison for quantitative results, measures of learning efficiency for interventions with some manner of computer assisted, or computer based instruction compared to traditional delivery methods are rare. Only two studies were identified as addressing this topic. The first was also reviewed in the learner attitude section. In their examination of knowledge gains, learning efficiency and learner satisfaction between an online tutorial program and printed materials, Bell et al. (2000) assessed students (n = 162) enrolled in family medicine and internal medicine residency programs at four universities. Results indicated no significant difference in posttest scores between those students using the online tutorial and the printed text materials. However, those utilizing the online tutorial spent less time studying (p< 0.001), demonstrating greater learning efficiency. The second study, also reviewed in the learner attitudes section was a simple survey study design to collect student attitudes regarding a distance medical course. In a self-assessment of knowledge gained, 18 of 32 students indicated they had acquired new knowledge, and 10 indicated that learning with the online text was more efficient than learning with a conventional textbook (Horsch et al. 2000). As with program cost, a fourth category offered by Chumley-Jones et al. (2002), this category of evaluative research regarding computer assisted, or computer based instruction requires further exploration.

Methodology

The present study examined the effectiveness of the HDD menu trainer in improving a student's ability to navigate and manipulate the MQ-9 menu structure, as well as potential impacts of either traditional, blended, or distance instruction on this process. Using the HDD menu trainer developed under the "Science and Technology for Warfighter Training and Aiding." Cooperative Agreement between AFRL and UND, pretests and posttests were used to measure learner knowledge gain. Learner attitude was assessed using a satisfaction survey.

Sample

The sample for this study consisted of individuals both with and without FAA pilot certification at the University of North Dakota John D. Odegard School of Aerospace Sciences (*n* = 15). Of this sample, 3 participants held no FAA pilot certificate, 5 participants held a Private Pilot Certificate, and 7 participants carried Commercial Pilot certification. The average subject age was 27.73. Subject responses were not separated by race or gender, and no subject's results were excluded from analysis. Participants were randomly assigned to one of three groups (i.e. Traditional, Distance, and Blended) receiving various instructional interventions with respect to MQ-9 HDD menus.

Instrument

The HDD menu trainer, developed by UND's Aerospace Network was designed to familiarize students with the layout and manipulation of the HDD menus for either the MQ-1 or MQ-9. The trainer contains (1) a tutorial describing menu layout, menu navigation, button types, and button arrangement, (2) a walk-through function, which guides students through each root menu and its submenus, (3) an exercise function, which tests the student's ability to navigate and execute specific commands within a set time limit, and finally (4) a freeplay function, which allows the students to navigate and explore the HDD menus without specific focus or limits on time.

The menu trainer was delivered to the distance and blended groups via an open source, online Learning Management System (LMS) administered by the researcher. All subjects had access to the LMS for completion of the pretest and posttest measures. Subjects were briefed on use of the LMS at the start of the intervention.

The pretest and posttest measures utilized a modified version of the HDD menu trainer's exercise function. These assessments, designed by an Original Equipment Manufacturer (OEM) certified MQ-9 IP, reflect those menu functions most commonly used or most critical for gauging a student's expertise with navigating and manipulating the HDD menus. Roughly 25 pilot orientated menu functions were selected from the pool of 260 which constitute the menu trainer's exercise function, and were adapted for delivery as the pretest and posttest measures. These measures, like the menu trainer's exercise function, measure the student's ability to navigate and execute specific commands within a set time limit. Performance was assessed according both the speed and accuracy of the student's response.

Data Collection and Analysis

This study was reviewed and approved by the University of North Dakota's Institutional Review Board. Subjects were informed of the study with advertisements posted throughout the campus aerospace facilities as well as the aviation student email listserve. Subjects were briefed on the purpose and nature of the study prior to participation. Due to the sensitive nature of the MQ-9 HDD menus, participants were also required to present proof of U.S. citizenship by means of a passport, and/or birth certificate and driver's license and sign an International Traffic in Arms Regulations (ITAR) Statement of Understanding.

Subjects were randomly assigned to one of three study groups to receive instruction on navigating and manipulating the HDD menus of the MQ-9. As illustrated in Table 1, students assigned to the distance group were granted access to the HDD menu trainer. Subjects assigned to the blended group were granted access to the HDD menu trainer, but also attended a classroom discussion guided by an Original Equipment Manufacturer (OEM) certified MQ-9 Instructor Pilot (IP). Subjects assigned to the traditional group were not granted access to the

HDD menu trainer, but received a lecture and lesson on the HDD Menus from an OEM certified MQ-9 IP. The lesson completed by the traditional group was conducted using an MQ-9 part-task trainer which simulated the same HDD menus but provided no innate instructional aspects (i.e. no tutorial, walk-through, or exercise functions). The layout and functionality of the menus simulated in this part-task trainer were identical to those used in the pretest and posttest measures, as well as those used by the distance and blended groups.

Table 1, Research Design

	Traditional Group	Blended Group	Distance Group
HDD Menu Trainer	No	Yes	Yes
MQ-9 Instructor Pilot	Yes	Yes	No

Descriptive and inferential statistics were collected from the data. The means, standard deviations, minimum, maximum, range, and measures of skewness and kurtosis indices were calculated using raw scores from each group. A one way ANOVA was used to assess differences between the groups on pretest, posttest, and percent change scores. In cases where parametric assumptions were violated, Kruskal-Wallis non-parametric procedures were used to assess potential relationships. Significance in all statistical tests were set at a minimum of $p<0.05$.

Results

Learner Knowledge Gains

Illustrated in Table 2 are descriptive statistics for each of the three groups in their pretest, posttest, and percent change measures. Each task in the parallel pretest and posttest measures was assigned 15 possible points. Points were deducted for incorrect keystrokes as well as when a task could not be completed inside 30 seconds. If a task was skipped, a score of 0 was assigned. Percent change was calculated as the difference between the pretest and posttest score divided by the pretest Score. Also included in Table 2 are z-scores for the skewness and kurtosis of each factor's score distribution. For these measures, absolute values greater than 1.96 indicate significantly non-normal distributions at $p<0.05$ (Field, 2009). Except for skewness in the percent change measure of the distance group, all measure distributions failed to differ significantly from a normal distribution in either skewness or kurtosis.

Table 2, Descriptive Results According to Instructional Method

	N	Mean	SD	Minimum	Maximum	Z skewness	Z kurtosis
PRE-TEST							
Traditional	5	244.00	68.58	157.00	324.00	0.00	-0.86
Blended	5	264.60	49.26	191.00	309.00	-0.87	-0.23
Distance	5	270.40	56.79	175.00	326.00	-1.70	1.57
POST-TEST							
Traditional	5	331.40	26.95	308.00	365.00	0.67	-1.45
Blended	5	334.00	27.59	299.00	371.00	0.21	-0.26
Distance	5	332.00	25.95	308.00	366.00	0.48	-0.58
PERCENT CHANGE							
Traditional	5	42.91	33.28	12.65	96.18	1.34	0.73
Blended	5	28.85	18.41	6.47	56.54	0.72	0.58
Distance	5	27.26	28.14	5.12	76.00	2.07*	1.92

* Indicates significance at the 0.05 level

Results of the one way ANOVA (Table 3) comparing pretest, posttest, and percent change scores between groups found no significant differences between the three groups on any of the measures. Although non-normality was noted in the skewness of the distance group in percent change, the same patterns of significance were noted using nonparametric Kruskal-

Wallis procedures comparing the mean ranks of percent change, as well as pretest and posttest scores, with respect to instructional method.

Table 3, One Way ANOVA Results According to Instructional Method

	Traditional Group		Blended Group		Distance Group		
	Mean	SD	Mean	SD	Mean	SD	P
PRE-TEST	244.00	68.58	264.60	49.26	270.40	56.79	0.761
POST-TEST	331.40	26.95	334.00	27.59	332.00	25.95	0.987
PERCENT CHANGE	42.91	33.28	28.85	18.41	27.26	28.14	0.620

* Indicates significance at the 0.05 level

A mixed ANOVA indicated the trainer significantly improved performance from pretest to posttest scores across all groups $F(1,12) = 49.01$ ($p<0.001$), however comparing these scores by instructional intervention (i.e. Traditional, Blended, and Distance) found no significant effect. To summarize, an overall effect of instruction was observed, but did not vary across the three types of instructional intervention.

Regarding pilot certification.

Analysis of pretest and posttest scores, as well as percent change in scores with respect to pilot certification revealed several relationships meriting consideration for future studies in this area. In Table 4, results of a one way ANOVA and Tukey *post hoc* analysis indicated that participants holding a commercial pilot certificate scored significantly higher on the pretest than those with no FAA pilot certification ($p<0.05$). No significant effect of pilot certification was found in an analysis of the post test scores. Furthermore, significantly higher percent changes from pretest to posttest were observed in participants with no FAA pilot certificate than those with commercial certificates ($p<0.05$). Again, a similar pattern of results were found when analysis was repeated using the Kruskal-Wallis procedure.

Table 4, One Way ANOVA results According to Pilot Certification

	None (n = 3)		Private (n = 5)		Commercial (n = 7)		
	Mean	SD	Mean	SD	Mean	SD	P
PRE-TEST	191.67	17.00	258.80	67.11	289.43	27.92	0.024*
POST-TEST	306.00	6.25	339.20	25.15	339.00	23.90	0.114
PERCENT CHANGE	60.45	14.01	37.88	34.05	17.77	10.47	0.041*

* Indicates significance at the 0.05 level

Learner Attitudes

A learner satisfaction survey was used to gauge participant satisfaction with the instruction they received. Participants were asked to respond to 8 statements regarding course satisfaction on a five point Likert scale (1=*Strongly Disagree*; 5=*Strongly Agree*). Sum totals and descriptive statistics for these responses are found in Table 5 below. While results of a one way ANOVA did not indicate a significant difference between course satisfaction and instructional method, patterns in the open ended responses offer some differentiation.

Table 5, Descriptive Results of Learner Attitude

	N	Mean	SD	Minimum	Maximum	Z skewness	Z kurtosis
ATTITUDE							
Traditional	5	29.20	5.45	22	35	-0.59	-0.99
Blended	5	32.60	6.23	23	39	-1.02	-0.38
Distance	5	29.00	1.00	28	30	0.00	-1.50

* Indicates significance at the 0.05 level

Open-ended responses to the prompts “Please describe improvements, if any, which would better assist your learning of the course material.” and “Please describe specific aspects of the course or instruction which promoted your learning." provide qualitative context. Members of the traditional group commonly felt that additional time and access to the HDD trainer would have better assisted their learning "… As someone who prefers to study alone, access to the trainer", "More time to teach the material", "More time with software" and "I would have benefitted from some practice exams at home." While the ability to govern instructional pace was a common theme in aspects of the course which promoted learning for members of the distance group, preference for an introductory lecture preceding self-study was noted as a way to better assist their learning. In the blended group, the combination of self-paced practice and the availability of instructor expertise in classroom discussions surfaced as positive aspects of the course.

Learning Efficiency

The traditional group was presented a 15 minute lecture followed by a simulated lesson in a part-task trainer Ground Control Station (GCS) for 45 minutes. As a single crew includes 1 pilot position and 1 sensor operator, this instruction only permitted 2 individuals to work directly with the IP at a time, while the remainder of the class observed. Following this lesson, participants were not allowed access to the menus excluding a 1 hour practice period in the simulated GCS. Self-reported study times for the distance group indicated an average of 1.3 hours of effort (0.84 *SD*) with the HDD menu trainer. Finally the self-reported study times for the blended group showed an average of 3.5 hours of effort (2.58 *SD*) preceding a 1 hour classroom discussion and review prior to the posttest.

Discussion

The results above demonstrate that the HDD menu trainer is effective in improving a student's ability to navigate and manipulate the MQ-9 menu structure. Results for learner knowledge gains, learner attitudes, and learning efficiency offer preliminary indications of the trainer’s potential as training aid in blended pedagogy, as well as standalone teaching tool in distance pedagogy. Similar to many previous efforts reviewed, the HDD menu trainer was at least as effective as the traditional method of instruction currently used in terms of learner knowledge gains. Although inferential results of the learner satisfaction survey did not reflect differing levels of satisfaction, written responses to the open ended portions of the instrument indicated that learners clearly identified with classic strengths and weaknesses of both traditional and distance pedagogies. The group receiving traditional instruction benefitted from the interaction and expertise of the live instructor, but requested additional time with the material or ways to study according to their individual needs. Members of the distance group, meanwhile, appreciated the ability to self-govern the pace of their learning but noted instructor availability as a way to improve their learning.

While it may have been anticipated that the blended group would outperform the other groups, benefitting from the advantages of instructor availability as well as the ability to govern their own preferences for pace and duration of instruction, the relatively small sample sizes likely affected this in two ways. First, if instructional method commands only a small effect size on learner knowledge gains, much larger sample sizes will be required to reliably detect a genuine effect when one exists. Second, as overall class size approaches the size of a single RPA crew, the unique differences between the instructor delivered portions of the blended and traditional approaches lessen. As class size approaches the size of a single crew, the lecture received by the traditional group increasingly resembles the individual attention normally reserved for individual lessons. Likewise, with fewer members of the blended group, individual members may benefit less from the questions and discussion generated between their peers and the instructor. As such, it may be that the blended pedagogy has a greater effect on learning knowledge gain and learning efficiency (in terms of instructor time) as class size increases.

Conclusion and Future Studies

As the availability and capability of instructional technologies continues to expand, opportunities to adapt, validate, and improve pedagogy accordingly are many. Extant literature reflecting evaluative efforts on distance and blended instruction generally report that these instructional methodologies are able to perform at least as well as traditional methods and in some circumstances, better. Blending the advantages of traditional face-to-face instruction with the benefits of computer aided delivery systems for learners is the focus of blended learning. The purpose of this pilot study has been to examine the expertise of students in navigating and manipulating the HDD menus of MALE RPA to assess (1) the effectiveness of the HDD menu trainer, and (2) its potential for use in traditional, blended, or distance instructional methods. Results of a mixed ANOVA indicated the trainer significantly improved performance from pretest to posttest scores across all groups ($p<0.001$), but comparisons by instructional intervention (i.e. Traditional, Blended, and Distance) found no significant effect. A lack of significant differences between pretest, posttest, and percent change scores between groups indicates that the HDD menu trainer may be assumed as equally effective in terms of learner knowledge gains across the instructional designs examined.

Exploration of the relationship between pilot certification and performance revealed an additional aspect influencing MALE RPA training, which must be controlled in future studies seeking variation uniquely attributable to instructional method. This pilot study found that learners holding a commercial pilot certificate scored significantly higher on the pretest than those with no FAA pilot certification ($p<0.05$). Such tendencies beg further investigations into the relationship of FAA pilot certification and MALE RPA training. What skills, knowledge, or experience, captured by these aviation benchmarks, accounts for the increased initial performance? Is the lack of significant difference between posttest scores with respect to certification the result of an artificial ceiling effect with the instrument? Does the ability to navigate and manipulate these menus represent understanding of their function? Perhaps considerations such as these can be used to adapt initial operations training in these platforms to the qualifications of those best qualified or most likely to be entering this new and rapidly evolving discipline.

As demand for MALE RPA pilots and sensor operators grows, adapting pedagogy and technologies to provide the highest standard of instruction at the greatest efficiency will remain an enormous challenge for all. Future studies involving the HDD menu trainer are underway utilizing the results of this pilot effort to isolate the unique variance in performance explained by instructional method and possible interactions between instruction and pilot certification. Informed by the results of this study, these efforts will utilize larger samples to map this relationship. Other studies are encouraged to document and reflect on learning efficiency, investigating whether use of such training aids can reduce instructor and/or simulator training time while engendering equivalent knowledge, skills, and abilities. Examining the pedagogy of MALE RPA training with consideration to learner knowledge gains, learner attitude, and learning efficiency will support the comprehensive understanding necessary to advance and mature this training domain.

References

Adler, M., & Johnson, K. (2000). Quantifying the literature of computer-aided instruction in medical education. *Academic Medicine 75 (10)*, 1025-1028.

Agius, R., & Bagnall, G. (1998). Development and evaluation of the use of the internet as an educational tool in occupational and environmental health and medicine. *Occupational Medicine, 48 (5)*, 337-343.

Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., & Burrell, N. (2004). Evaluating the effectiveness of distance learning: a comparison using meta-analysis. *Journal of Communication, 54*, 402-420.

Baumlin, K., Bessette, M., Lewis, C., & Richardson, L. (2000). EMCyberSchool: An evaluation of computer-assisted instruction on the internet. *Academic Emergency Medicine, 7 (8)*, 959-962.

Bedwell, W., & Salas, E. (2010). Computer-based training: capitalizing on lessons learned. *International Journal of Training and Development 14 (3)*, 239-249.

Bell, D., Fonarow, G., Hays, R., & Mangione, C. (2000). Self-study from web-based and printed guidline materials: A randomized, controlled trial among resident physicians. *Annals of Internal Medicine, 132 (12),* 938-946.

Block, A., Felix, M., Udermann, B., Reineke, D., & Murray, S. (2008). Achievement and satisfaction in an online versus a traditional health and wellness course. *Journal of Online Learning and Teaching, 4 (1),* 57-66.

Boyle, T., Bradley, C., Chalk, P., Jones, R., & Pickard, P. (2003). Using blended learning to improve student success rates in learning to program. *Journal of Educational Media 28 (2&3),* 165-178.

Burke, M., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R., & Islam, G. (2006). Relative Effectiveness of Worker Safety and Health Training Methods. *American Journal of Public Health, 96 (2)*, 315-324.

Chumley-Jones, H., Dobbie, A., & Alford, C. (2002). Web-based learning: sound educational method or hype? A review of the evaluative literature. *Academic Medicine, 77*, 86-93.

Cohen, V. (2003). A model for assessing distance learning instruction. *Journal of Computing in Higher Education, 14 (2)*, 98-120.

Curran, V., Hoekman, T., Gulliver, W., Landells, I., & Hatcher, L. (2000). WWeb-based continuing medical education (II): Evaluation study of computer-mediated continuing medical education. *The Journal of Continuing Education in the Health Profession, 20*, 106-119.

DeBourgh, G. (1999). Technology is the tool, teaching is the task: Student satisfaction in distance learning. *Society for Information Technology and Teacher Education International Conference*, (pp. 1-8). San Antonio.

DeRouin, R., Fritzsche, B., & Salas, E. (2004). Optimizing e-learning: reserach-based guidelines for learner-controlled training. *Human Resource Management Journal, 43 (2&3)*, 147-162.

Engel, S., Crandall, J., Basch, C., Zybert, P., & Wylie-Rosett, J. (1997). Computer-assisted diabetes nutrition education increases knowledge and self-efficacy of medical students. *The Diabetes Educator, 23 (5)*, 545-549.

Field, A. (2009). *Discovering Statistics Using SPSS.* London: Sage Publications.

Francis, B., Mauriello, S., Phillips, C., Englebardt, S., & Grayden, S. (2000). Assessment of online continuing dental education in north carolina. *The Journal of Continuing Education in the Health Profession, 20*, 76-84.

Harris, J., Salasche, S., & Harris, R. (2001). Can internet-based continuing medical education improve physicians' skin cancer knowledge and skills? *Journal of General Internal Medicine, 16*, 50-56.

Horsch, A., Balbach, T., Melnitzki, S., & Knauth, J. (2000). Learning tumor diagnostics and medical image processing via the WWW - The case-based radiological textbook ODITEB. *International Journal of Medical Informatics, 58-59*, 39-50.

Hsu, L.-L., & Hsieh, S.-I. (2011). Effects of a blended learning module on self-reported learning performances in baccalaureate nursing students. *Journal of Advanced Nursing* , 2435-2444.

Kronz, J., Silberman, M., Allsbrook Jr., W., & Epstein, J. (2000). A web-based tutorial improves practicing pathologists' gleason grading of images of prostate carcinoma specimens obtained by needle biopsy: Validation of a new medical education paradigm. *The American Cancer Society, 89 (8)*, 1818-1823.

Lipman, A., Sade, R., Glotzbach, A., Lancaster, C., & Marshall, M. (2001). The incremental value of internet-based instruction as an adjunct to classroom instruction: A prospective randomized study. *Academy of Medicine, 76 (10)*, 72-76.

Melton, B., Graf, H., & Chopak-Foss, J. (2009). Achievement and satisfaction in blended learning versus traditional general health course designs. *International Journal for the Scholarship of Teaching and Learning, 3 (1)*, 1-13.

Osguthorpe, R., & Graham, C. (2003). Blended learning environments: Definitions and directions. *The Quarterly Review of Distance Education, 4 (3)*, 227-233.

Overbaugh, R. (1994). Reserach-based guidelines for computer-based instruction development. *Journal of Reserach on Computing in Education 27 (1)*, 29-48.

Pereira, J., Pleguezuelos, E., Meri, A., Molina-Ros, A., Molina-Tomas, C., & Masdeu, C. (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical Education, 41*, 189-195.

Perryer, G., Walmsley, A., Barclay, C., Shaw, L., & Smith, A. (2002). Development and evaluation of a stand-alone web-based CAL program. *European Journal of Dental Education, 4 (3)*, 118-123.

Rivera, J., & Rice, M. (2002). A comparison of student outcomes & satisfaction between traditional & web based course offerings. *Online Journal of Distance Learning Administration, 5 (3)*.

Rose, M., Frisby, A., Hamlin, M., & Jones, S. (2000). Evaluation of the effectiveness of a web-based graduate epidemiology course. *Computers in Nursing, 18 (4)*, 162-167.

Sakowski, H., Rich, E., & Turner, P. (2001). Web-based case simulations for a primary care clerkship. *Academic Medicine, 76 (5)*, 547.

Smyth, S., Houghton, C., Cooney, A., & Casey, D. (2012). Sutdents' experiences of blended learning across a range of postgraduate programmes. *Nurse Education Today, 32*, 464-468.

Woo, M., & Kimmick, J. (2000). Comparison of internet versus lecture instructional methods for teaching nursing research. *Journal of Professional Nursing, 16 (3)*, 132-139.

Study II
Waller and Stupnisky (Manuscript)

Medium Altitude Long Endurance RPA Training: Evaluating Blended Learning

Rights
This study is under review for publication by the International Journal of Aviation, Aeronautics, and Aerospace (IJAAA).

Abstract

The Heads Down Display (HDD) Menu Trainer – a stand-alone software trainer – was developed to familiarize students in Remotely Piloted Aircraft training with the layout and manipulation of the HDD menus for either the MQ-1 or MQ-9. Preliminary work by Waller et al. (2016) established the efficacy of the HDD Menu Trainer in improving student performance from pretest to posttest scores across several modalities (i.e. traditional, blended, and distance). Recognizing that students holding pilot certification scored higher in some aspects of the HDD Menu Trainer, this study sampled students across a curriculum to assess whether performance with the HDD Menu Trainer would differ across modalities (i.e. traditional, blended, and distance) when FAA pilot certification was controlled.

Results of a mixed factorial ANCOVA indicated the effectiveness of the HDD menu trainer once more through a main within-subjects effect of performance and performance was again higher for students holding an FAA pilot certificate than for those without. However, modality failed to demonstrate a significant interaction effect with student performance from pretest to posttest. These results affirm that even outside the variation which should be attributed to a student’s pilot certification, the HDD Menu Trainer demonstrates equal effectiveness when used in blended and distance modalities. These results support several prior works finding blended learning applications to be at least as effective as other modalities.

As blended, flipped, and hybrid learning models are increasingly expected within higher education curriculums, future work is anticipated in the construct of student engagement (Borup et al., 2020; Halverson & Graham, 2019).

Medium Altitude Long Endurance RPA Training: Evaluating Blended Learning

In 2003, Osguthorpe and Graham situated their understanding of blended learning according to aspects of both modality (i.e. the mode of delivery) and pedagogy (i.e. the method of teaching). Since then, assessments of blended learning at the course-level have established its effectiveness through comparison to traditional models – commonly construed as face-to-face (Porter et al., 2014; Waller et al., 2016).

Statement of the problem.

Between 2011 and 2017 the “Science and Technology for Warfighter Training and Aiding.” Cooperative Agreement between the University of North Dakota and the Air Force Research Laboratory (AFRL) produced curriculum for Medium Altitude, Long Endurance Remotely Piloted Aircraft (MALE RPA) pilots and sensor operators. From these efforts was developed a Heads Down Display (HDD) Menu Trainer as a stand-alone software trainer to familiarize students with the layout and manipulation of the HDD menus for either the MQ-1 or MQ-9.

Preliminary work by Waller et al. (2016) established the efficacy of this HDD menu trainer in improving student performance from pretest to posttest scores across several modalities (i.e. traditional, blended, and distance). Waller et al. also noted that participants with greater levels of Federal Aviation Administration (FAA) pilot certification scored significantly higher on the pretest measure of the HDD Menu Trainer but lacked a sufficient sample to assess FAA pilot certification as a covariate.

Data collection across the curriculum, rather than within a course, was needed to assess whether student performance across modalities would begin to differ when the model allowed FAA pilot certification to covary

Purpose of the study.

The purpose of this study was to sample students across a curriculum, rather than within a course, to assess whether student performance with the HDD Menu Trainer would differ across modalities (i.e. traditional, blended, and distance) when FAA pilot certification was allowed to covary in the analysis.

Literature Review

Measures such as (1) student evaluations and satisfaction (Horsch et al., 2000; Hsu & Hsieh, 2011; Smyth et al., 2012), (2) student performance and achievement, (Allen et al., 2004; Baumlin et al., 2000; Bell et al., 2000; Block et al., 2008; Boyle et al., 2003; Curran et al., 2000; Engel et al., 1997; Francis et al., 2000; Harris et al., 2001; Kronz et al., 2000; Lipman et al., 2001; Melton et al., 2009; Perryer et al., 2002; Rivera & Rice, 2002; Rose et al., 2000; Sakowski et al., 2001; Woo & Kimmick, 2000), the Sloan-C Pillars (Laumakis et al., 2009), and even the confidence of students (Pereira et al., 2007) have all seen use in situating instructional models (e.g. face-to-face, blended, and online) according to modality.

As the adoption of blended learning progressed, proponents predicted it would become the ‘new normal’ within higher education (Norberg et al., 2011). Accepting the course-level effectiveness of blended learning, the sections below review institutions and administrations seeking a better understanding of how blended learning might be strategically implemented at scale.

University of Granada, Spain.

Among the first examples aggregating data across curriculums is a blended learning initiative evaluated by Lopez-Perez et al. (2011) at the University of Granada, Spain. First year undergraduate students (*n* = 985) – enrolled in Business Administration and Management, Economics Business Studies, and the Business Administration/Law courses – provided their perceptions of the courses via a 13-item survey (Lopez-Perez et al., 2011). The students' performance was also measured by (1) the proportion of students sitting the final exam (the 'non-dropout rate') and (2) the proportion of passing grades (Lopez-Perez et al., 2011).

Results indicated that blended learning reduced dropout rates and increased exam passing rates (Lopez-Perez et al., 2011). A comparison of regression models indicates that students' motivation during the face-to-face portion of their blended course were predictive of their final grade ($p< 0.01$), over and above the variation explained by their age, gender, average grade prior to entering the course, and attendance (Lopez-Perez et al., 2011). Lopez-Perez et al. (2011) offer that the motivation, satisfaction, and perceived utility of blended learning may influence student performance in an indirect way.

University of Central Florida, United States.

Moskal et al. (2013) assess the performance of blended learning efforts at the University of Central Florida (UCF). With an interest in improving teaching and informing institutional policymaking, Moskal et al. investigated how student satisfaction, success, and withdrawal related to course modality (i.e. blended, fully online, face-to-face, blended lecture capture, and lecture capture). Course ratings from academic years 2008 to 2011 were indexed by modality (Moskal et al., 2013).

A large sampling (n = 913,688) of student satisfaction reflected "… the blended modality [enjoyed] the highest percentage (52%) of "excellent" responses producing a 4% marginal advantage over online and face-to-face courses that [were] tied at 48%..." (Moskal et al., 2013, p. 19). From this finding, the university used regression tree analysis to identify aspects of the instructor and course which lead to an overall rating of 'excellent' (Moskal et al., 2013).

The analysis of Moskal et al. (2013) found that if UCF students rated the instructor's (1) ability to facilitate learning, (2) communication skill, and (3) respect and concern for students as ‘excellent’, the probability of the course receiving an overall rating of ‘excellent’ moved to .97 – regardless of the modality. Encouraged by this finding, Moskal et al. conducted a hierarchical logistic regression indicating that over and above the predictive power of demographic characteristics associated with students, the addition of these three instructor qualities is able to increase R^2 by 0.719. Regardless of modality, which does not enter the model, these three items are proposed as high-impact areas for improving pedagogy (Moskal et al., 2013).

When student rates of success – measured as earning a passing grade – and withdrawal were evaluated against modality, courses in the blended learning category yielded the highest success rates of 90.8% and saw withdrawal at roughly half the rate (2.8%) of lecture capture courses (5.3%) (Moskal et al., 2013).

York University, Canada.

At York University, Owston et al. (2013) examined the relationship between student perceptions and achievement in blended learning courses. Following a multi-year initiative to increase blended learning, students (n = 577) were surveyed from eleven (11) blended learning courses. In an Analysis of Covariance (ANCOVA) model, responses to a 31-item survey were entered as the independent variables, cumulative grade point averages (CGPA) were entered as a covariate, and final grade for the blended coursework was entered as the dependent measure of achievement (Owston et al., 2013).

Results indicated higher achievement (i.e. a final grade) for students who strongly agreed with the statements ‘I am satisfied with this [blended] course’ and ‘I would take another blended course’ – $F(4,448) = 12.69$, $p = .000$, $\eta^2 = .102$ and $F(5,447) = 6.30$, $p = .000$, $\eta^2 = .066$, respectively, with the estimated marginal mean of final grades corrected for CGPA. Owston et al. (2013) conclude, "… that the highest achievers were most satisfied with their blended course, would take one again, and preferred the blended format over fully face-to-face or online [courses]" (p. 41). The opposite was found for low achieving students.

Further results from the ANCOVA model indicated that high achieving students found that blended learning offered (1) convenience, and (2) reduced travel time and expenses – $F(5,445) = 6.37$, $p = .000$, $\eta^2 = .067$ and $F(5,443) = 5.56$, $p = .000$, $\eta^2 = .059$, respectively (Owston et al., 2013). When assessing the relationship between engagement in blended learning and achievement, the largest effect was found in responses to the statement asking whether students were engaged more in their current blended course than other face-to-face courses they had taken, $F(5,444) = 15.99$, $p = .000$, $\eta^2 = .153$ (Owston et al., 2013). All but one of the twelve Likert statements related to engagement indicated significant differences between high and low achievers. For the inquiry related to students' perceptions of learning, Owston et al. (2013) relay a significant relationship between responses to the statement 'Compared to typical face-to-face courses I have taken… this course has improved my understanding of key concepts', $F(5,446)= 6.38$, $p= .000$, $\eta^2 = .067$.

Following York University's implementation of a major blended learning initiative, Owston et al. (2013) offer, "high achievers are very satisfied with the blended format, find blended learning to be convenient and flexible, are very engaged in their studies, and appear to learn key concepts better" (p. 43). The endorsement supports the university's interests with the caveat from Owston et al. that blended courses may not be as suitable for low achievers.

While several of the higher-education efforts above were funded internally, some noted grant support from the NGLC awarded jointly to the American Association of State Colleges and

Universities (AASCU) and the University of Central Florida (UCF) (Porter et al., 2014), or a Sloan fluency/localness grant awarded to the University of Wisconsin–Milwaukee (UWM) (Graham et al., 2013).

Methodology

The present study examined the impact of modality (i.e. traditional, blended, or distance) in learning the HDD menus of a MALE RPA while controlling for FAA pilot certification. Using the HDD Menu Trainer developed under the “Science and Technology for Warfighter Training and Aiding.” Cooperative Agreement between the Air Force Research Laboratory and the University of North Dakota, pretests and posttests were used to measure learner knowledge gain.

Sample.

The sample for this study consisted of individuals both with and without FAA pilot certification at the University of North Dakota John D. Odegard School of Aerospace Sciences (n=102). Of this sample, 26 participants held no FAA pilot certificate, 48 participants held a Private Pilot certificate, and 27 participants carried Commercial Pilot certification. Average age was 22.93 (SD=5.68). Participants were assigned to modality groups (i.e. Traditional, Distance, and Blended) by class, with each class receiving various instructional interventions for teaching the Heads Down Display (HDD) Menus of the MQ-9.

Instrument.

The HDD Menu Trainer, developed by UND, was designed to familiarize students with the layout and manipulation of the HDD menus for either the General Atomics MQ-1 or MQ-9. The trainer contains (1) a tutorial describing menu layout, menu navigation, button types, and button arrangement, (2) a walk-through function, which guides students through each root menu and its submenus, (3) an exercise function, which tests the student’s ability to navigate and execute specific commands within a set time limit, and finally (4) a freeplay function, which allows students to navigate and explore the HDD menus without specific focus or limits on time.

The menu trainer was delivered to the distance and blended groups via an open source, online Learning Management System (LMS). All participants had access to the LMS for completion of the pretest and posttest measures. Participants were briefed on use of the LMS at the start of the intervention.

The pretest and posttest measures utilized a modified version of the HDD Menu Trainer's exercise function. Designed by an Original Equipment Manufacturer (OEM) certified MQ-9 instructor pilot, these assessments represented those menu functions most commonly used or most critical for gauging a student's expertise with navigating and manipulating the HDD menus. Roughly 25 pilot orientated menu functions were selected for the pretest and posttest from the trainer's 260 exercise functions, and were adapted for delivery as the pretest and posttest measures. As with the trainer's exercise function, the student's ability to navigate and execute specific commands within a set time limit were assessed. Performance was measured according both the speed and accuracy of the student's response.

Data collection and analysis.

This study was reviewed and approved by the applicable Institutional Review Board. Participants were informed of the study with advertisements posted throughout the campus aerospace facilities as well as the aviation student email listserve. Participants were briefed on the purpose and nature of the study prior to participation. Due to the sensitive nature of the MQ-9 HDD Menus, participants were also required to present proof of U.S. citizenship by means of a passport, and/or birth certificate and driver's license and sign an International Traffic in Arms Regulations (ITAR) Statement of Understanding.

The variety of modalities examined in this study were largely delivered during existing aviation courses, and random assignment among the groups should not be assumed. Preliminary work has indicated that pilot certification significantly affects pretest performance. To mitigate possible effects of this stratified sampling, participant level of FAA pilot certification has been controlled wherever learner knowledge gains are assessed across pedagogies.

Each modality group received instruction on navigating and manipulating the HDD menus of the MQ-9. Illustrated in Table 1 below, students of the distance group were only granted access to the HDD Menu Trainer. Students of the blended group were granted access to the HDD Menu Trainer, but also attended a classroom discussion guided by an OEM certified MQ-9 Instructor Pilot (IP). Students assigned to the traditional group were not granted access to the HDD Menu Trainer, but rather received a lecture and simulator lesson on the HDD menus from an OEM certified MQ-9 IP. To ensure the same menu structure was represented in the instruction of the Traditional group and the pretest and posttest measures, the freeplay function of the HDD Menu Trainer was utilized in the simulated lesson. The version of the HDD Menu Trainer provided for this purpose had only freeplay functionality, the tutorial, walk-through, and exercise functions were disabled.

Table 6, Research Design

	Traditional Group	Blended Group	Distance Group
HDD Menu Trainer	Freeplay Only	Full	Full
MQ-9 Instructor Pilot	Yes	Yes	No

Results

Illustrated in Table 2 are descriptive statistics for each of the three groups in their pretest, posttest, and percent change measures. Each task in the parallel pretest and posttest measures was assigned 15 possible points. Points were deducted for incorrect keystrokes as well as when a task could not be completed inside 30 seconds. If a task was skipped, a score of 0 was assigned. Percent change was calculated as the difference between the pretest and posttest score divided by the pretest score.

While significant departures from normality were noted among each of the pretest, posttest, and percent change distributions in the Kolmogorov-Smirnov and Shapiro-Wilk tests, the *F* statistic has been found to be robust against such violations. Because parametric assumptions may not be considered tenable, the results of the inferential procedures that follow should be interpreted with caution.

Table 7,
Descriptive Statistics for Student Performance

	N	Mean	SD	Minimum	Maximum
PRE-TEST					
Traditional	39	203.95	69.47	63.00	324.00
Blended	29	210.80	60.72	103.00	311.00
Distance	30	235.24	70.86	14.00	326.00
POST-TEST					
Traditional	39	271.26	71.98	45.00	365.00
Blended	29	289.40	45.30	195.00	371.00
Distance	30	287.62	72.84	13.00	373.00
PERCENT CHANGE					
Traditional	39	43.60	48.44	-75.41	183.05
Blended	29	48.42	48.57	-15.67	192.23
Distance	30	25.98	30.19	-15.63	111.39

* Indicates significance at the .05 level
** Indicates significance at the .01 level
*** Indicates significance at the .001 level

An independent samples *t*-test (see Table 3) was used to compare the hours of self-study reported by students of the distance (*M* = 1.25, *SD* = 1.00) and blended (*M* = 1.22, *SD* = 1.51) modalities. Students in both of these groups had remote access to the HDD menu trainer, while members of the traditional group did not. Results indicated no difference in amounts of self-study between students in the blended and distance groups $t(54)=-0.08$, *ns*.

Table 8,
Comparison of Self-Reported Hours of Study

	n	*M (SD)*	Mean Difference	*t*	*df*
Modality					
Blended	30	1.22 (1.52)	-0.03	-0.08	54
Distance	26	1.25 (1.00)			

* Indicates significance at the .05 level
** Indicates significance at the .01 level
*** Indicates significance at the .001 level

Performance across Modality

Results of a mixed factorial ANCOVA analyzed variation unique to modality (i.e. Traditional, Distance, and Blended) while controlling for whether or not a student held an FAA pilot certificate. Results, shown in Table 4, indicated the effectiveness of the HDD menu trainer once more through a main within-subjects effect of performance. That is, regardless of modality, and controlling for pilot certification, posttest scores were higher than pretest scores. As shown in the estimated marginal means plotted in Figure 1, modality failed to demonstrate a significant interaction effect with student performance from pretest to posttest.

Although the same mixed factorial ANCOVA procedure indicated a significant between-group main effect of pilot certification, no interactive effect was noted between student performance with the HDD Menu Trainer and FAA pilot certification. Regardless of modality, student performance was again higher for students holding an FAA pilot certificate than for those without.

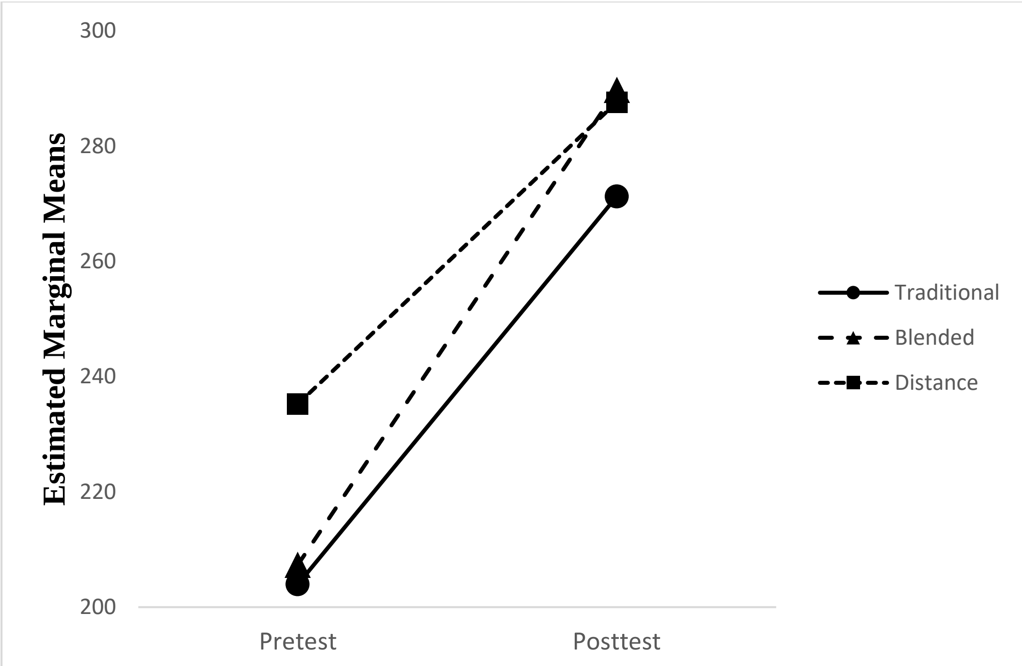


Figure 1,
Estimated Marginal Mean Performance by Modality

Table 4,
Regressing Performance across Modality (Pilot Certification Controlled)

	df	MS	*F*	η^2
Performance	1	45212.63	27.65***	.23
Performance * Modality	2	3396.18	2.07	.04
Performance * Pilot Certification (Covariate)	1	299.83	0.18	.00
Error (Performance)	93	15203.25		
Modality	2	3896.21	0.56	.01
Pilot Certification	1	27550.44	3.97*	.04
Error	93	6939.72		

* Indicates significance at the .05 level
** Indicates significance at the .01 level
*** Indicates significance at the .001 level

Discussion and Conclusion

This study sampled students across a curriculum to assess whether student performance with the HDD Menu Trainer would differ across modalities (i.e. traditional, blended, and distance) when FAA pilot certification was controlled. Waller et al. (2016) noted that students holding pilot certification scored higher in some aspects of the HDD Menu Trainer. Here, whether a student holds an FAA pilot certification is entered as a covariate to control for these differences and better isolate variation which may be attributed to modality. Once again, the HDD Menu Trainer demonstrates (1) an ability to improve student ability in navigating and manipulating the HDD menus for the MQ-9 and (2) a significant between-subjects main effect on performance for students holding an FAA pilot certificate. Neither pilot certification nor modality was found to have a significant interactive effect on student performance.

Prior work assessing blended learning applications has spanned several countries and disciplines. Like many of these works (Allen et al., 2004; Baumlin et al., 2000; Bell et al., 2000; Block et al., 2008; Boyle et al., 2003; Curran et al., 2000; Engel et al., 1997; Francis et al., 2000; Harris et al., 2001; Kronz et al., 2000; Lipman et al., 2001; Melton et al., 2009; Perryer et al., 2002; Rivera & Rice, 2002; Rose et al., 2000; Sakowski et al., 2001; Woo & Kimmick, 2000), this study compared modalities using student performance and achievement. Like many of these, this study found its blended learning application to be at least as effective as other modalities.

Lopez-Perez et al. (2011) utilized several regression models to better isolate the effect of motivation during the face-to-face portion of a blended learning experience, and Moskal et al. (2013) utilized a hierarchical logistic regression to explain the effect of three instructor qualities – over and above the predictive power of students' demographic characteristics. As Owston et al. (2013) would enter cumulative grade point averages as an ANCOVA model covariate, so this study sought to increase the sensitivity of its model by designating a covariate of its own related to student performance. The ANCOVA results above affirm that even outside the variation which should be attributed to a student's pilot certification, the HDD Menu Trainer demonstrates equal effectiveness when used in blended and distance modalities.

Future Directions

Blended learning has long been situated in terms of both modality and pedagogy (Osguthorpe & Graham, 2003). As the blended learning model undergoes ongoing assessment and increasing integration within higher education, interests have begun to pivot toward goals such as (1) enhancing pedagogy and increasing access (Graham et al., 2005), (2) more efficient use of classroom resources and extending campus outreach (Graham et al., 2005; Moskal et al., 2013), or even (3) adapting the educational paradigm for "… the 'new type of learner' enrolling at the university" (Carbonell et al., 2013, p. 32).

Having so reviewed strategic integration of instruction which "… combines face-to-face with distance delivery systems…" (Osguthorpe & Graham, 2003, p. 227), a brief treatment of transitions to technology-assisted instruction which have not been strategic is also warranted on behalf of educational technology and instructional design scholars. The term 'emergency remote teaching' has recently emerged as a way to distinguish the mandatory transition that many institutions implemented to prevent the spread of the virus that causes COVID-19 (Hodges et al., 2020). Where modality alone would closely associate the emergency remote teaching of Hodges et al. (2020) or the HyFlex model explained by Irving (2020) with blended learning, proponents are already separating the three on pedagogical terms (Saichaie, 2020).

Although discussion – or perhaps more accurately – clarification surrounding modality has resurged with emergency remote teaching, the future directions of inquiry specific to blended learning appear to be focusing increasingly on the student engagement (Borup et al., 2020; Halverson & Graham, 2019). The study of this construct – its measurement and supporting mechanisms – are well situated as blended, flipped, and hybrid learning models are increasingly expected within higher education curriculum all around the globe (Saichaie, 2020).

References

Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., & Burrell, N. (2004). Evaluating the effectiveness of distance learning: a comparison using meta-analysis. *Journal of Communication*, *54*, 402–420.

Baumlin, K., Bessette, M., Lewis, C., & Richardson, L. (2000). EMCyberSchool: An evaluation of computer-assisted instruction on the internet. *Academic Emergency Medicine*, *7*, 959–962.

Bell, D., Fonarow, G., Hays, R., & Mangione, C. (2000). Self-study from web-based and printed guidline materials: A randomized controlled trial among resident physicians. *Annals of Internal Medicine*, *132*, 938–946.

Block, A., Felix, M., Udermann, B., Reineke, D., & Murray, S. (2008). Achievement and satisfaction in an online versus a traditional health and wellness course. *Journal of Online Learning and Teaching*, *4*, 57–66.

Boyle, T., Bradley, C., Chalk, P., Jones, R., & Pickard, P. (2003). *Using blended learning to improve student success rates in learning to program*. Journal of Educational Media 28 (2&3).

Carbonell, K. B., Dailey-Hebert, A., & Gihselaers. (2013). Unleashing the creative potential of faculty to create blended learning. *Internet and Higher Education*, *18*, 29–37.

Curran, V., Hoekman, T., Gulliver, W., Landells, I., & Hatcher, L. (2000). Web-based continuing medical education (II): Evaluation study of computer-mediated continuing medical education. *The Journal of Continuing Education in the Health Profession*, *20*, 106–

119.

Engel, S., Crandall, J., Basch, C., Zybert, P., & Wylie-Rosett, J. (1997). Computer-assisted diabetes nutrition education increases knowledge and self-efficacy of medical students. *The Diabetes Educator*, *23*, 545–549.

Francis, B., Mauriello, S., Phillips, C., Englebardt, S., & Grayden, S. (2000). Assessment of online continuing dental education in north carolina. *The Journal of Continuing Education in the Health Profession*, *20*, 76–84.

Graham, C., Allen, S., & Ure, D. (2005). Benefits and Challenges of Blended Learning Environments. In *M* (pp. 253–259). IGI Global.

Graham, C., Woodfield, W., & Harrison, B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *Internet and Higher Education*, *18*, 4–14.

Harris, J., Salasche, S., & Harris, R. (2001). Can internet-based continuing medical education improve physicians' skin cancer knowledge and skills? *Journal of General Internal Medicine*, *16*, 50–56.

Horsch, A., Balbach, T., Melnitzki, S., & Knauth, J. (2000). Learning tumor diagnostics and medical image processing via the WWW - The case-based radiological textbook ODITEB. *International Journal of Medical Informatics*, *58–59*, 39–50.

Hsu, L.-L., & Hsieh, S.-I. (2011). Effects of a blended learning module on self-reported learning performances in baccalaureate nursing students. *Journal of Advanced Nursing*, 2435–2444.

Irvine, V. (2020). The Landscape of Merging Modalities. *EDUCAUSE Review*, *55*(4).

Kronz, J., Silberman, M., Jr., A., W., & Epstein, J. (2000). A web-based tutorial improves practicing pathologists' gleason grading of images of prostate carcinoma specimens obtained by needle biopsy: Validation of a new medical education paradigm. *The American Cancer Society*, *89*(8), 1818–1823.

Laumakis, M., Graham, C., & Dziuban, C. (2009). The Sloan-C Pillars and Boundary Objects As a Framework for Evaluating Blended Learning. *Journal of Asynchronous Learning Networks*, *13*(1), 75–87.

Lipman, A., Sade, R., Glotzbach, A., Lancaster, C., & Marshall, M. (2001). The incremental value of internet-based instruction as an adjunct to classroom instruction: A prospective randomized study. *Academy of Medicine*, *76*(10), 72–76.

Lopez-Perez, M., Perez-Lopez, M., & Rodriguez-Ariza, L. (2011). Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers and Education*, *56*, 818–826.

Melton, B., Graf, H., & Chopak-Foss, J. (2009). Achievement and satisfaction in blended learning versus traditional general health course designs. *International Journal for the Scholarship of Teaching and Learning*, *3*, 1–13.

Moskal, P., Dziuban, C., & Hartman, J. (2013). Blended learning: A dangerous idea? *Internet and Higher Education*, *18*, 15–23.

Norberg, A., Dziuban, C., & Moskal, P. (2011). A Time Based Blended Learning Model. *On the Horizon*, *19*(3), 207–216. https://doi.org/http://dx.doi.org/10.1108/10748121111163913

Osguthorpe, R., & Graham, C. (2003). Blended learning environments: Definitions and directions. *The Quarterly Review of Distance Education*, *4*(3), 227–233.

Owston, R., York, D., & Murtha, S. (2013). Student perceptions and achievement in a university blended learning strategic initiative. *Internet and Higher Education*, *18*, 38–46.

Pereira, J., Pleguezuelos, E., Meri, A., Molina-Ros, A., Molina-Tomas, C., & Masdeu, C. (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical Education*, *41*, 189–195.

Perryer, G., Walmsley, A., Barclay, C., Shaw, L., & Smith, A. (2002). Development and evaluation of a stand-alone web-based CAL program. *European Journal of Dental Education*, *4*(3), 118–123.

Porter, W., Graham, C., Spring, K., & Welch, K. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers and Education*, *75*, 185–195.

Rivera, J., & Rice, M. (2002). A comparison of student outcomes & satisfaction between traditional & web based course offerings. *Online Journal of Distance Learning Administration*, *5*.

Rose, M., Frisby, A., Hamlin, M., & Jones, S. (2000). Evaluation of the effectiveness of a web-based graduate epidemiology course. *Computers in Nursing*, *18*, 162–167.

Sakowski, H., Rich, E., & Turner, P. (2001). Web-based case simulations for a primary care clerkship. *Academic Medicine*, *76*(547.).

Smyth, S., Houghton, C., Cooney, A., & Casey, D. (2012). Sutdents' experiences of blended learning across a range of postgraduate programmes. *Nurse Education Today*, *32*, 464–468.

Waller, Z., Petros, T., Bridewell, J., Kroeber, S., & Nowatzki, N. (2016). Medium Altitude Long Endurance Remotely Piloted Aircraft Training: A Pilot Study in Blended Learning. *Journal of Unmanned Aerial Systems*, *1*, 6–16. https://doi.org/10.31356/avi-fac0002

Woo, M., & Kimmick, J. (2000). Comparison of internet versus lecture instructional methods for teaching nursing research. *Journal of Professional Nursing*, *16*, 132–139.

Study III
Waller and Petros (Manuscript)

The Effectiveness of Operator State Monitoring
in Measuring RPA Training

Rights
This study was presented to a national audience and is under review for publication by the Collegiate Aviation Review International (CARI).

Waller, Z. P. (Presenter & Author), Petros, T. (Co-Author), 2021 Collegiate Aviation Education Conference & Expo, "Cognitive State Measurement in Remotely Piloted Aircraft Training," University Aviation Association, Memphis, TN. (October 9, 2021).

Abstract

Electroencephalograms (EEG) have been shown to reflect workload levels and sustained attention during training and learning, however a limited number of studies has examined the effectiveness of EEG in operational settings (Bernhardt et al., 2019; Mathan & Yeung, 2015; Mills et al., 2017; Yuan et al., 2014). The purpose of this research is to determine whether EEG technology is sensitive to changes in the cognitive workload and task engagement of remote pilots during simulated training events with the MQ-1. EEG data was collected from remote pilots (n = 10) during simulated MQ-1 RPA training events in the PRINCE device.

Estimates of the Advanced Brain Monitoring, Inc. (ABM) cognitive metrics for high engagement and workload were averaged for the duration of the checklist as well as each leg of the flight pattern. Results of one-way repeated measures ANOVAs showed that the cognitive state metric for engagement $F(11,8704)=4.87$, $p<0.001$ and workload $F(11,8328)=10.03$, $p<0.001$ varied significantly within the flight pattern. Results of a paired sample t-test $t(8348)=14.21$, $p<.001$ indicated that workload was significantly lower ($M=0.5536$, $SD=0.16$) during legs of the flight pattern assisted by the heading hold function of the autopilot than those legs where remote pilots were unassisted by this automation ($M=0.5718$, $SD=0.16$).

As with prior works in operational aviation settings, EEG-based cognitive state metrics demonstrated an ability to detect subtle changes in operator workload (Aricò et al., 2016; Bernhardt et al., 2019; Borghini et al., 2015). The NASA TLX was administered to collect a subjective measure of workload but no significant association was observed between the subjective and EEG-based measures of workload.

The Effectiveness of Operator State Monitoring in Measuring RPA Training

Cognitive workload and task engagement are common constructs in human performance research, and represent the supply-demand relationship of cognitive resources and the attentional resources available to attend a task, respectively (Bernhardt et al., 2019). Over the past two decades, equipment and indices have been developed to measure these constructs of performance in laboratory settings using basic cognitive tasks (Berka et al., 2007; Johnson et al., 2011). These cognitive state measures have been proposed for assessing the effectiveness of training and simulation programs because they are able to assess change which is not obvious from task performance alone (Berka et al., 2007; Parasuraman, 2015). Electroencephalograms (EEG) have been shown to reflect workload levels and sustained attention during training and learning, however a limited number of studies has examined the effectiveness of EEG in operational settings (Bernhardt et al., 2019; Mathan & Yeung, 2015; Mills et al., 2017; Yuan et al., 2014).

Purpose of the Study

The purpose of this research is to determine whether EEG technology is sensitive to changes in the cognitive workload and task engagement of remote pilots during simulated training events with the MQ-1.

Literature Review

The construct of mental or cognitive workload has donned several definitions. Recent works have approached the construct as "… the dynamic relationship between the resources that are needed to carry out a task and the ability of the operator to adequately supply those resources." (Bernhardt et al., 2019, p. 83) Throughout these definitions a few points have become common, In general, mental workload theory assumes that:

(a) people have limited cognitive and attentional capacity,

(b) different tasks will require different amounts of processing resources, and

(c) two individuals might be able to perform a given task equally well, but differently in terms of brain activation. (Aricò et al., 2016, p. 299)

In contrast to cognitive workload, the construct of engagement has been described as "… the availability of attentional resources and the mobilization of resources for efficient processing of task-related stimuli…" (Bernhardt et al., 2019, p. 83) Rather than the supply and demand relationship of cognitive workload, engagement is typically associated with states such as concentration or sustained attention (Bernhardt et al., 2019).

By merit of its economy in cost and size, as well as its resolution in spatial and temporal terms, EEG has settled into a role of providing neurophysiological measurement of cognitive processes (Mills et al., 2017). EEG measures the minute voltage that passes through the scalp as the result of coordinated firing of billions of neurons in the brain (Mathan & Yeung, 2015). Because these voltages are viewed simultaneously across several regions of the brain, machine-learning techniques are employed to characterize the unique patterns of neural response to cognitive effort (Mathan & Yeung, 2015). Measuring these voltages at several locations, multivariate discriminate functions characterize these patterns and may or may not be calibrated on an individual basis with baseline tasks. All of these methods ultimately use EEG signals to produce a single-dimensional estimate of effort (Mathan & Yeung, 2015). Patterns of neural activity and EEG features have been correlated in laboratory settings with the constructs of workload and engagement.

Development of EEG measures in Laboratory Settings

A wide variety of methods, algorithms, and models have been leveraged to index the construct of cognitive workload from EEG measurements, however the brain's theta (4–8 Hz) and alpha (8–12 Hz) activity from the prefrontal cortex (PFC) and the posterior parietal cortex (PPC) regions consistently contribute to the analysis (Aricò et al., 2016). Aricò et al. (2016) summarize several prior works to offer that the theta frequency band of the PFC is typically positively correlated with cognitive workload, while the alpha frequency of the PPC is typically negatively correlated with cognitive workload (Gevins & Smith, 2000). Significant contributions from the delta, beta, and gamma frequency bands appear to be less common (Aricò et al., 2016). Development of the workload metric used here – the posterior probabilities of high and low workload commercially available through Advanced Brain Monitoring, Inc. (ABM) – reaches back to the laboratory tasks of Berka et al. (2007). In this effort, Berka et al. validated their metric with data acquired as participants (*n* = 13) completed five laboratory tasks developed by Lockheed Martin. The tasks included between three and six levels of difficulty and were performed in the following order (1) grid, (2) forward digit span, (3) mental arithmetic, (4) backward digit span, and (5) trails (Berka et al., 2007). At each level of difficulty, participants were surveyed for a subjective measure of workload on three 100-point scales. The questions were (1) "How much mental energy did you exert on this task level?", (2) "Objectively, how difficult was this task level?", and (3) "How much attention did you focus on this task level?" (Berka et al., 2007, p. B233).

The finished workload classifier of Berka et al. (2007) utilized EEG signals from the C3-C4, Cz-PO, F3-Cz, Fz-C3, and Fz-PO sites. Thirty (30) EEG features were used to calculate a workload metric which was significantly higher during the encoding period of the Verbal Paired

Associate test (VPA), image-learning, and memory tests (Berka et al., 2007). The workload metric also correlated with subjective metrics and increased linearly with difficulty in the forward and backward-digit-span, grid-recall, and mental-addition tests (Berka et al., 2007). Berka et al. and Johnson et al. (2011), would also develop and validate four measurements of engagement. These engagement metrics utilize twenty-three (23) EEG features from the Fz-POz and Cz-POz sites and are calibrated according to a 3-choice vigilance task, a visual psychomotor vigilance task, and an auditory psychomotor vigilance task. These tasks (i.e. benchmarks) are performed by each participant allowing ABM's model to produce posterior probabilities of (1) high engagement, (2) low engagement, (3) relaxed wakefulness, and (4) sleep onset on a scale ranging from 0.00 to 1.00 (Berka et al., 2007).

In 2014, Sciarini et al. would find ABM's workload metric to be sensitive to the changes in cognitive effort involved in completion of a Stroop task. Sciarini et al. (2014) explain, "The Stroop effect is elicited in experiments by manipulating the text of the name of a color, for example 'brown.' The stimulus is manipulated by presenting the text in the same color or in a different color than brown so that there is either congruence or incongruence between text and the color…" (p. 216). Longer reaction times associated with incongruence are attributed to a disruption in attentional allocation (Sciarini et al., 2014).

Moving beyond the laboratory setting, EEG-based metrics of workload and engagement are now seeing application and assessment in operational settings (Borghini et al., 2012; Marcel & Millán, 2007; Schubert et al., 2008; Venthur et al., 2010; Welke et al., 2009), within education and training (Mathan & Yeung, 2015; Mills et al., 2017; Yuan et al., 2014), and also aviation.

Application of EEG Measurement in Aviation Settings

In 2020, Belkhiria and Peysakhovich published a review of efforts involving both EEG and Electrooculogram (EOG) within the field of aeronautics between 2010 and 2020. The purpose of this review was – in part – "… to provide methodological guidelines for beginners and experts when applying [combined EEG and EOG] in environments outside the laboratory, with a particular focus on human factors and aeronautics." (Belkhiria & Peysakhovich, 2020, p. 1). Although the number of participants in each study varies, the majority of the reviewed datasets contained fewer than 30 participants.

As early as 2010, the alpha frequency recorded just prior to stimulus was found to be a promising metric for active monitoring of both engagement as well as workload (Baldwin et al., 2010). In a visual search task which simulated the role of a Unmanned Aerial Vehicle (UAV) operator, Baldwin et al. (2010) offered that the alpha frequency demonstrated potential "… as an index of when an operator may be more error prone or when a learner may be reaching a state where he or she is less likely to benefit from an instructional strategy." (p. 9)

In 2014, Borghini et al. began investigating the application of cognitive metrics in the evaluation of Air Traffic Control (ATC) students (n = 6) learning a new Air Traffic Management (ATM) task. Not only EEG, but also Electrocardiograms (ECG) and EOG signals were collected from participants. Results from a one-way repeated measures ANOVA indicated that theta frequencies over the frontal cortex (i.e. AF3, AF4, F3, Fz, and F4) varied significantly as training progressed across one week (Borghini et al., 2014). A second repeated measures ANOVA showed that alpha frequencies over the parietal cortex (i.e. P3, Pz and P4) decreased as these ATC students progressed through training (Borghini et al., 2014).

Borghini et al. (2015) leveraged a Mental Workload Index (MWL) to assess the impact of a variety of avionic technologies on the cognitive workload of helicopter pilots. Eight (8) EEG channels were collected and the MWL index was "… defined as the ratio between the frontal theta and parietal alpha EEG [Power Spectral Density] PSD values." (Borghini et al., 2015, p. 6182) The study was designed around simulated operations during which participants used technologies which included a Head-Up Display (HUD), a Head-Mounted Display (HMD) and a Synthetic Vision System (SVS) (Borghini et al., 2015). Results indicated that the workload index was lower when pilots used the HUD as opposed to all other technologies.

In 2016, Aricò et al. measured the cognitive workload of French Air Traffic Controllers (n = 12) during simulated ATM scenarios. In this effort, a machine-learning algorithm was used to index workload from EEG signals at eight (8) sites – positioned at Fz, F3, F4, AF3, AF4, Pz, P3, and P4 of the 10-20 standard. The EEG measure of workload was correlated against a subjective measure of workload as well as examined for reliability across one month (Aricò et al., 2016).

The experimental scenarios were accomplished across 45 minutes and varied in complexity (i.e. 'EASY', 'MEDIUM', and 'HARD'). Workload demands during the scenario were adjusted by varying the number of aircraft, the number and type of clearances required, and the number and trajectory of interfering flights. Two pseudo pilots also interacted with the participants to simulate real-flight communications. The controllers were presented each of the three levels in 15 minute increments and a random order and provided subjective ratings of their workload on a five-point scale every 3 minutes – ranging from 1 'very easy' to 5 'very difficult' (Aricò et al., 2016).

Results of a one-way ANOVA on the Aricò et al. (2016) EEG measure of workload indicated that the controllers' cognitive workload changed significantly during the simulated scenario. Controllers' workload for EASY was significantly lower than MEDIUM, that MEDUIM was significantly lower than HARD, and finally, that workload for EASY was significantly lower than HARD (Aricò et al., 2016).

Aricò et al. (2016) also provided a contrast between workload measures produced by EEG data and subjective measures. A Pearson's correlation coefficient demonstrated a pronounced and positive relationship between the EEG-based workload measure and subjective measure provided by the controllers ($r = 0.856$, $p = 0.0002$) and the expert observers ($r = 0.797$, $p = 0.0011$).

Bernhardt et al. (2019) measured workload and engagement among ATC students (n = 47) with varying levels of experience in a simulated ATC environment. Results of a 2 (experience) by 5 (difficulty) mixed factorial design – with experience as the between-subjects factor and scenario phase difficulty as the within-subjects factor – indicated that less experienced controllers exhibited higher engagement than more experienced controllers. Although ABM's metric for average probability of workload was sensitive to changes in workload throughout the scenario, it did not differentiate between experience groups. While pupil diameter was anticipated to correlate with ABM's workload metric – when averaged across the five phases of the scenario – the two measures were not correlated $r(43) = -0.25$, $p = 0.098$ (Bernhardt et al., 2019). Bernhardt et al. posit that the workload construct may not be a unitary, or that the pupil diameter measurement may include multiple physiological responses beyond workload alone – such as alertness or engagement.

The present study provides one of the first attempts to employ ABM's operator state monitoring to measure workload and task engagement during simulated training events of remote pilots.

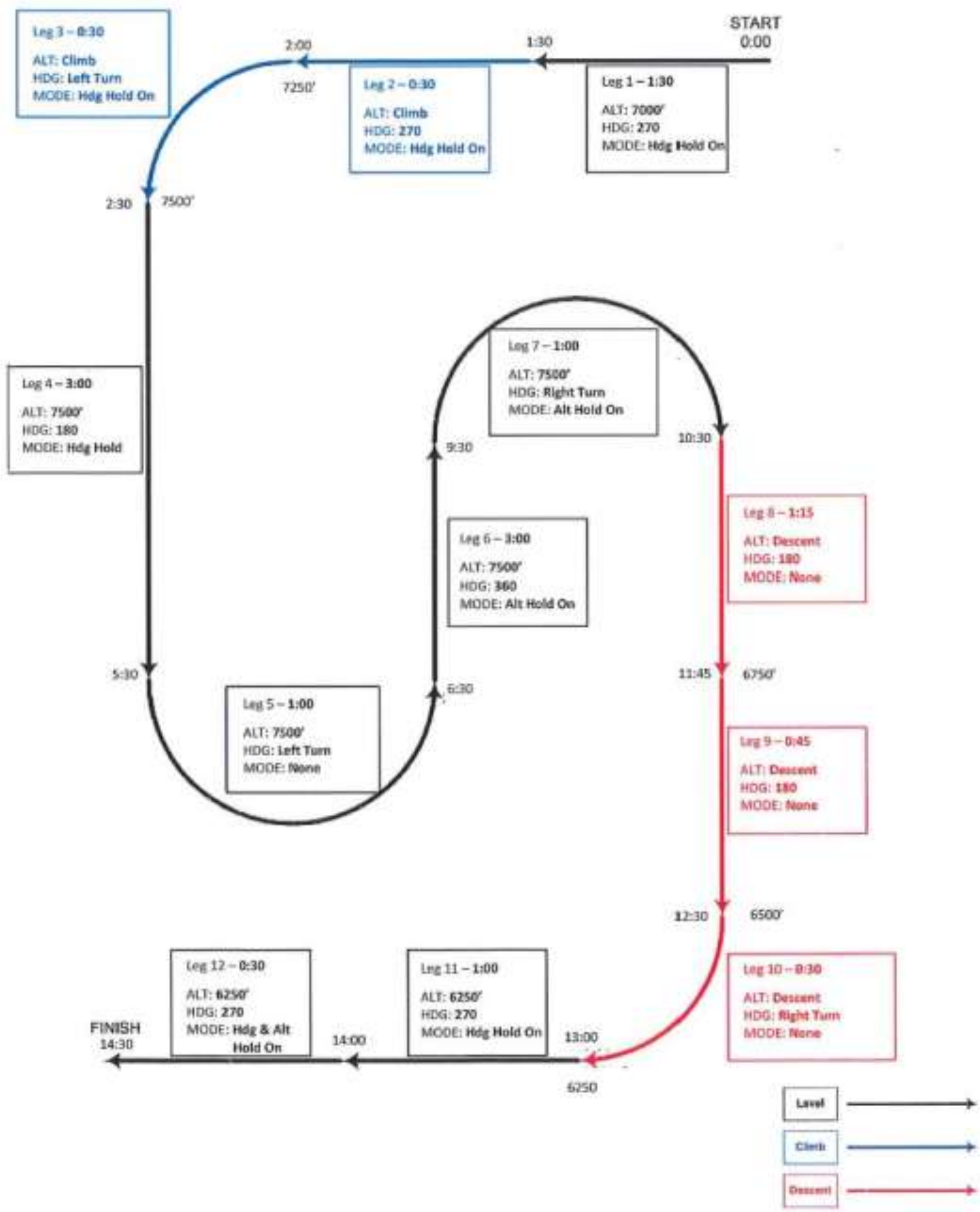
Methodology

This study was reviewed and approved by the applicable Institutional Review Board. Participants were briefed on the purpose and nature of the study prior to participation. All participants were enrolled in training curriculum, which requires proof of U.S. citizenship by means of a passport, and/or birth certificate and driver's license as well as completion of an International Traffic in Arms Regulations (ITAR) Statement of Understanding.

Sample, Instruments, and Data Collection

EEG data was collected from remote pilots (n = 10) during simulated MQ-1 RPA training events in the PRINCE device. Posterior probabilities of ABM's high workload and engagement metrics were collected throughout the simulation. The lesson calls for approximately 1.2 hours of contact time with the remote pilot, a checklist, and a flight pattern with 12 distinct legs (see Figure 1). At this point in their curriculum, the student flight crew – one remote pilot and one sensor operator – are expected to complete lesson tasks with instructor guidance. To begin this course of training, participants had earned a commercial pilot certificate with instrument ratings and accrued approximately 200 hours of total flight time in manned aircraft.

Figure 2, Simulated Flight Pattern



Prior to the lesson, participant were administered a questionnaire of their flight experience, completed the Vocabulary subtest from the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997), as well as the Vandenberg and Kuse Mental Rotation Test (Vandenberg & Kuse, 1978). Once these measures were complete, EEG signals were collected using the ABM B-Alert X-24 wireless Bluetooth system. The B-Alert X-24 incorporates 20 electrodes placed at the Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, POz, and O2 sites of the international 10/20 system. To establish benchmarks for each participant in cognitive load, drowsiness and distractibility the B-Alert X-24 uses three baseline cognitive assessment tasks (1) a three choice vigilance task, (2) a visual stimulus response task, and (3) an eyes closed, auditory stimulus-response task (Advanced Brain Monitoring, 2009).

ABM produces its cognitive workload metric using two models – one produced using a forward digit span (FDS) task and the second using a backward digit span (BDS) task. Both models produce probabilities ranging from 0.00 to 1.00 with those closer to 1.00 reflecting higher workload. In the interest of generalizability, the cognitive state metric used to measure workload in this study was the mean between the FDS and BDS models – also produced by ABM (Advanced Brain Monitoring, 2009)

Probabilities of cognitive states – such as high and low engagement, cognitive workload, distraction, and sleep onset – were calculated by ABM metrics (Berka et al., 2007; Johnson et al., 2011). Probabilities from each 1 second (i.e., epoch) of the simulated lesson were generated for each cognitive state – ranging from 0.00 to 1.00 (Advanced Brain Monitoring, 2009). The start and end of the checklist and each flight leg were manually marked in the recording by a research technician to facilitate analysis of the EEG-based metrics.

Results

From the sample of remote pilots, the average probabilities of ABM's high engagement and high workload metrics are provided in Table 1. Analysis and visualizations were produced using the R language and RStudio software (RStudio Team, 2020).

Table 9, Probability of High Engagement and Workload

	N (epochs)	Engagement Mean	SD	Workload Mean	SD
CHECKLIST	10 (16,939)	0.52	.40	0.57	.16
FLIGHT PATTERN					
Leg 1 (90 sec)	10 (897)	0.47	.40	0.57	.15
Leg 2 (30 sec)	10 (304)	0.53	.40	0.56	.15
Leg 3 (30 sec)	10 (301)	0.52	.42	0.57	.14
Leg 4 (180 sec)	10 (1,866)	0.50	.40	0.54	.16
Leg 5 (60 sec)	10 (622)	0.49	.39	0.58	.15
Leg 6 (180 sec)	10 (1,810)	0.45	.40	0.56	.16
Leg 7 (60 sec)	10 (587)	0.49	.41	0.60	.15
Leg 8 (75 sec)	10 (789)	0.50	.40	0.56	.15
Leg 9 (45 sec)	10 (456)	0.43	.40	0.56	.16
Leg 10 (30 sec)	10 (309)	0.43	.41	0.59	.15
Leg 11 (60 sec)	10 (607)	0.50	.40	0.56	.16
Leg 12 (30 sec)	10 (308)	0.42	.40	0.55	.15

Estimates of ABM's cognitive metrics for high engagement and workload were averaged for the duration of the checklist as well as each leg of the flight pattern, producing Figures 2 and 3 below. Results of one-way repeated measures ANOVAs showed that the cognitive state metric for engagement, $F(11,8704)=4.87$, $p<0.001$, and workload $F(11,8328)=10.03$, $p<0.001$ varied significantly within the flight pattern. Results of a paired sample *t*-test, $t(8348)=14.21$, $p<.001$, indicated that workload was significantly lower ($M=0.5536$, $SD=0.16$) during legs of the flight pattern assisted by the heading hold function of the autopilot than those legs where remote pilots were unassisted by this automation ($M=0.5718$, $SD=0.16$).

Figure 3, Cognitive State Metric for Workload: ABM's High Workload Metric during Checklist and Flight Pattern Events

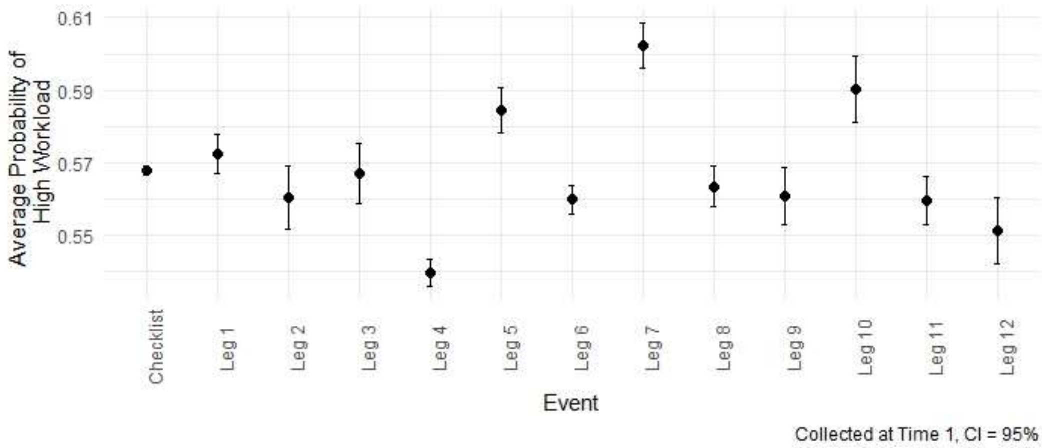
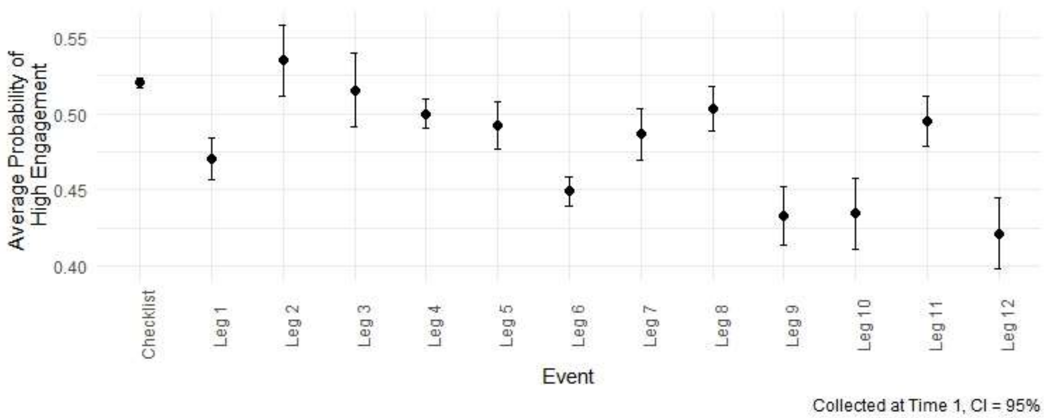


Figure 4, Cognitive State Metric for Engagement: ABM's High Engagement Metric during Checklist and Flight Pattern Events



Each remote pilot provided a subjective evaluation of his or her workload following both the checklist and flight pattern tasks using the NASA TLX (Hart & Staveland, 1988). ABM's probability of high workload was averaged for each participant using each epochs which occurred during the checklist and flight pattern. A Pearson's product moment correlation was performed between the self-reported workload on the TLX and the ABM calculated cognitive state metric for high workload. Results indicated no significant relationship between the two measures during either the checklist $r = -0.02$, $t(8) = -0.05$, $p = 0.96$ or flight pattern events $r = -$

0.53, $t(8) = -1.79$, $p = 0.11$. When both events were taken together, the same result was found $r = -0.29$, $t(18) = -1.31$, $p = 0.21$.

Discussion and Conclusions

This study demonstrates that EEG technology - developed for cognitive state estimation in operational settings – is sensitive to changes in the cognitive workload and task engagement of remote pilots during simulated training events with the MQ-1. As with prior works in operational aviation settings, EEG-based cognitive state metrics demonstrated an ability to detect subtle changes in operator workload (Aricò et al., 2016; Bernhardt et al., 2019; Borghini et al., 2015). The NASA TLX was administered to collect a subjective measure of workload, but no significant association was observed between the subjective and EEG-based measures of workload.

The NASA TLX is collected in several prior studies. Borghini et al. (2014), for example, administered the NASA TLX as an alternate measure of workload where the instrument demonstrated how perceived workload significantly decreases as training in ATM tasks progresses within one week. However, the measure is rarely correlated directly with EEG-based metrics of workload. The absence of this procedure, in fact, is acknowledged as an explicit limitation in the work of Bernhardt et al. (2019).

The lack of association in our results contrasts with other reports of a positive correlation between the NASA TLX and workload (Mathan & Yeung, 2015). It also contrasts with the prominent positive association noted by Aricò et al. (2016) between a subjective workload measure of their own – referred to as the Instantaneous Self-Assessment (ISA) – and an EEG-derived measure.

The NASA TLX is noted as a hallmark among self-report measures but differs in important ways from both the ISA administered by Aricò et al. (2016) and the 6-point Likert scale of Mills et al. (2017). Even when administered using the iOS application, the NASA TLX requires operators to select between 15 pairwise comparisons which may contribute to the workload of the task before rating their effort on six scales. Whereas the pairwise weighing and multiple scales create a robust foundation for the NASA TLX, selecting a response ranging from 1 'very easy' to 5 'very difficult' on the ISA or responding on a simple 6-point Likert scale allows reporting with little interruption and while the task is underway. The three minute frequency of Aricò et al. for sampling subjective workload seems better suited to establishing associations with EEG-based metrics.

Here, EEG-based metrics for measuring cognitive states demonstrate a sensitivity for detecting variation during the training of RPA pilots. These results support design of a within-subjects methodology using EEG data to assess the effectiveness of RPA training over time.

Future Directions

The measures and methods of cognitive states are maturing and coalescing into fields such as augmented cognition and adaptive automation. Many applications within aviation have emphasized the potential of cognitive state monitoring in the interest of safety, acknowledging "... errors could arise from aberrant mental processes, such as inattention, poor motivation, loss of vigilance, mental overload, and fatigue, that negatively affect the user's performance" (Aricò et al., 2016, p. 296).

In terms of technology advancement and integration, future work might also be expected in the validation of systems employing (1) dry electrodes – which eliminate the conductive gel or saline patch required to reduce the skin's contact impedance – and (2) in-ear EEG devices –

which are simpler to place and diminish motion artifacts in the data (Belkhiria & Peysakhovich, 2020). Both of these technological adaptations would ease collection of EEG signals in operational environments and enable assessment in increasingly ecological environments.

Methodological and statistical advances are only increasing access to the potential of physiological signals like EEG. Belkhiria and Peysakhovich (2020) summarize,

> … advances in signal processing analysis provide a powerful tool for modeling complex probability distributions by automatically discovering intermediate abstractions from a huge amount of basic features. Deep machine learning and artificial intelligence have shown great promise in helping make sense of EEG signals… (p. 18)

As technological and statistical advancement expand collection and access to EEG-based cognitive state metrics, still other works have sought to relate mental workload – measured not by EEG but by either ECG or EOG – with other measures of human performance such as perceptual load, stress, and performance on a modified Fitts' task (Causse et al., 2016; Mallat et al., 2020; Mandrick et al., 2016). Still further effort may be anticipated here as augmented cognition is pursued as a method for optimizing learning performance (Mathan & Yeung, 2015), or predicting human error (Baldwin et al., 2010; Mazaheri et al., 2009), or even enabling single-pilot or reduced-crew operations in aviation (Schmid & Stanton, 2020). These distinct physiological metrics of workload could stand to further validate or augment the EEG-derived cognitive states employed above.

References

Advanced Brain Monitoring. (2009). *B-alert Live: User Manual*.

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Pozzi, S., & Babiloni, F. (2016). A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks. *Progress in Brain Research*, *228*, 295–328.

Baldwin, C., Coyne, J. T., Roberts, D. M., Barrow, J. H., Cole, A., Sibley, C., Taylor, B., & Buzzell, G. (2010). *Prestimulus alpha as a precursor to errors in a UAV target orientation detection task.* Naval Research Lab.

Belkhiria, C., & Peysakhovich, V. (2020). Electro-Encephalography and Electro-Oculography in Aeronautics: A Review Over the Last Decade (2010–2020). *Frontiers in Neuroergonomics*, *1*, 1–25. https://doi.org/doi: 10.3389/fnrgo.2020.606719

Berka, C., J., L. D., N., L. M., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5), B231–B244.

Bernhardt, K. A., Poltavski, D., Petros, T., Ferraro, F. R., Jorgenson, T., Carlson, C., & Iseminger, C. (2019). The effects of dynamic workload and experience on commercially available EEG cognitive state metrics in a high-fidelity air traffic control environment. *Applied Ergonomics*, *77*, 83–91.

Borghini, G., Arico`, P., Di Flumeri, G., Salinari, S., Colosimo, A., Bonelli, S., Napoletano, L., Ferreira, A., & Babiloni, F. (2015). Avionic technology testing by using a cognitive neurometric index: A study with professional helicopter pilots. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6182–6185. https://doi.org/http://dx.doi.org/10.1109/EMBC.2015.7319804

Borghini, G., Aricò, P., Ferri, F., Graziani, I., Pozzi, S., Napoletano, L., Imbert, J. P., Granger, G., Benhacene, R., & Babiloni, F. (2014). A neurophysiological training evaluation metric for air traffic management. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3005–3008. https://doi.org/http://dx.doi.org/10.1109/EMBC.2014.6944255

Borghini, G., Vecchiato, G., Toppi, J., Astolfi, L., Maglione, A., Isabella, R., Caltagirone, C., Kong, W., Wei, D., Zhou, Z., Polidori, L., Vitiello, S., & Babiloni, F. (2012). Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6442–6445. https://doi.org/http://dx.doi.org/10.1109/EMBC.2012.6347469

Causse, M., Imbert, J. P., Giraudet, L., Jouffrais, C., & Tremblay, S. (2016). The role of cognitive and perceptual loads in inattentional deafness. *Frontiers in Human Neuroscience*, *10*, 1–12. https://doi.org/10.3389/fnhum.2016.00344

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, *10*(9), 829–839.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139–183. https://doi.org/https://doi.org/10. 1016/S0166-4115(08)62386-9

Johnson, R. R., Popovic, D. P., Olmstead, R. E., Stikic, M., Levendowski, D. J., & Berka, C. (2011). Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological Psychology*, *87*(2), 241–250. https://doi.org/https://doi.org/10.1016/j.biopsycho.2011.03.003

Mallat, C., Cegarra, J., Calmettes, C., & Capa, R. L. (2020). A curvilinear effect of mental workload on mental effort and behavioral adaptability: an approach with the pre-ejection period. *Human Factors*, *62*(6), 928–939. https://doi.org/10.1177/0018720819855919

Mandrick, K., Peysakhovich, V., Rémy, F., Lepron, E., & Causse, M. (2016). Neural and psychophysiological correlates of human performance under stress and high mental workload. *Biological Psychology*, *121*, 62–73.

Marcel, S., & Millán, J. D. R. (2007). Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 743–752. https://doi.org/http:// dx.doi.org/10.1109/TPAMI.2007.1012

Mathan, S., & Yeung, N. (2015). *Extending the diagnostic capabilities of artificial intelligence-based instructional systems.* 51–60.

Mazaheri, A., Nieuwenhuis, I. L., Van Dijk, H., & Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Human Brain Mapping*, *30*(6), 1791–

1800.

Mills, C., Fridman, I., Soussou, W., Waghray, D., Olney, A. M., & D’Mello, S. K. (2017). Put your thinking cap on: detecting cognitive load using EEG during learning. *Seventh International Learning Analytics & Knowledge Conference*, 80–89.

Parasuraman, R. (2015). Neuroergonomic perspectives on human systems integration: Mental workload, vigilance, adaptive automation, and training. In D. A. Boehm-Davis, F. T. Durso, & J. D. Lee (Eds.), *APA handbook of human systems integration* (pp. 163–176). American Psychological Association.

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. http://www.rstudio.com/

Schmid, D., & Stanton, N. A. (2020). Progressing Toward Airliners’ Reduced-Crew Operations: A Systematic Literature Review. *The International Journal of Aerospace Psychology*, *30*(1–2), 1–24. https://doi.org/10.1080/24721840.2019.1696196

Schubert, R., Tangermann, M., Haufe, S., Sannelli, C., Simon, M., Schmidt, E. A., Kincsesc, W. E., & Curio, G. (2008). Parieto-occipital alpha power indexes distraction during simulated car driving. *International Journal of Psychophysiology*, *69*(3), 214. https://doi.org/http://dx.doi.org/10.1016/j.ijpsycho.2008.05.033

Sciarini, L. W., Grubb, J. D., & Fatolitis, P. G. (2014). Cognitive state assessment: Examination of EEG-based measures on a stroop task. *Human Factors and Ergonomics Society Annual Meeting*, 215–219.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604. https://doi.org/https://doi.org/10.2466/ pms.1978.47.2.599

Venthur, B., Blankertz, B., Gugler, M. F., & Curio, G. (2010). Novel applications of BCI technology: psychophysiological optimization of working conditions in industry. *2010 IEEE International Conference on Systems, Man and Cybernetics*, 417–421. https://doi.org/10.1109/ICSMC.2010.5641772

Wechsler, D. (1997). WAIS III: Wechsler adult intelligence scale. *Psychological Corporation*.

Welke, S., Jurgensohn, T., & Roetting, M. (2009). Single-trial detection of cognitive processes for increasing traffic safety. *International Technical Conference on the Enhanced Safety of Vehicles*, 1–10.

Yuan, Y., Chang, K. M., Taylor, J. N., & Mostow, J. (2014). Toward unobtrusive measurement of reading comprehension using low-cost EEG. *Fourth International Conference on Learning Analytics and Knowledge*, 54–58.

Discussion & Conclusions

Remote pilots bear a high cost for flight training; costs which can be reduced by minimizing contact time in flight training devices. Studies I through III assessed adaptations in method (i.e. trainers and modalities) and measures of performance (e.g. workload and task engagement) to improve the efficiency and quality of remote pilot training in simulated MQ-1 operations. This chapter reviews and reflects upon the program of research assembled above, presents the implications of these works for RPA training, acknowledges limitations, and anticipates future directions for research.

Review of Research Program and Implications

In the case of Study I and II, it was observed that students have difficulty navigating and manipulating a particular menu during their RPA training with the MQ-1 and MQ-9. The HDD Menu Trainer was developed to familiarize students with the layout and manipulation of these HDD menus but was unproven as a standalone trainer. Twenty-five (25) tasks in the HDD menus were selected by an OEM certified instructor pilot to measure student performance and the effectiveness of the trainer was demonstrated with a pretest/posttest design across modalities in Study I (Waller et al., 2016).

Recognizing that students holding pilot certification scored higher in some aspects of the HDD Menu Trainer, Study II sampled students across a curriculum to assess whether performance with the HDD Menu Trainer would differ across modalities (i.e. traditional, blended, and distance) when FAA pilot certification was controlled. Multiple regression models have been leveraged previously within higher education as efforts sought the unique variability attributable to blended learning. Methodologically, Study II is most similar to the ANCOVA procedures of Owston et al. (2013) which allowed cumulative grade point averages of students to covary as their model predicted performance in an academic course from student perceptions of blended learning.

Where modality failed to demonstrate an interaction effect with student performance from pretest to posttest in Study II, the ANCOVA results affirmed that even outside the variation which should be attributed to a student's pilot certification, the HDD Menu Trainer was equally effective when used in blended and distance modalities. It improves students' RPA training today as a standalone method of familiarizing students with the layout and manipulation of the HDD menus prior to beginning fifteen hours of simulated MQ-1 operations; a course of training which carries an average cost of $4,151 per student (John D. Odegard School of Aerospace Sciences, 2021).

Study III further expanded the examination of RPA training in simulated MQ-1 operations by assessing whether EEG technology would be sensitive to changes in the cognitive workload and task engagement of remote pilots once they reach simulated training events with the MQ-1. The B-Alert X-24 wireless Bluetooth system of ABM was used to collect EEG signals across 20 electrodes. The cognitive workload metric – utilizing thirty (30) EEG features from the C3-C4, Cz-PO, F3-Cz, Fz-C3, and Fz-PO sites – and the engagement metric – utilizing twenty-three (23) EEG features from the Fz-POz and Cz-POz sites and three benchmark tasks – were initially developed and validated by Berka et al. (2007) and Johnson et al. (2011) using laboratory tasks.

Moving from the assessment of RPA training methods, the design of Study III investigates cognitive state metrics as a novel measure of performance and is supported by similar assessments in operational aviation settings (Aricò et al., 2016; Belkhiria & Peysakhovich, 2020; Bernhardt et al., 2019; Borghini et al., 2014, 2015). The results of Study III demonstrate that the workload and engagement metrics produced by ABM are sensitive to changes in the cognitive states of remote pilots during simulated training events with the MQ-1. Variation in workload is greatest across 12 legs of a flight pattern and significantly reduced workload is specifically noted during legs of this flight pattern when remote pilots were assisted by the heading hold function of the autopilot. Whereas student completion standards depend today upon the nature of task completion (e.g. whether a task is completed 'with flight instructor guidance', 'with little flight instructor guidance', or 'without flight instructor guidance'), the results of Study III improve RPA training by demonstrating an alternate measure of competency or performance which may not be obvious from task completion alone.

Limitations

Studies I through III note a number of limitations. Although the pattern of significance and direction were consistent between the examination of HDD Menu Trainer performance of Studies I and II, skew and kurtosis within the performance (i.e., both pretest and posttest scores) likely contributed to significantly non-normal distributions. Though the *F* statistic has been found to be robust against such violations (Howell, 2016), parametric assumptions may not be tenable and the results should be interpreted with caution.

The results of Study III found no association between the self-reported workload of the NASA TLX and the cognitive state metric for high workload calculated by ABM. While this lack of association between cognitive state metrics and the NASA TLX – a hallmark among self-report measures of workload – sacrifices some level of criterion validity, correlations between this instrument and alternate measures of workload or difficulty have seen mixed results (Borghini et al., 2014, 2015; Mathan & Yeung, 2015). It is possible that the design of the Study III tasks may have obscured the strength of the association.

It is noted above that ABM's cognitive workload metric varied significantly across legs of the flight pattern, some of which were as short as 30 seconds in duration. To avoid interrupting the progress of the flight pattern, the NASA TLX – which requires participants to select between 15 pairwise comparisons and rate their effort across six scales – was only administered once the entire flight pattern was completed. Criterion validity for ABM's cognitive state metrics may be pursued with more success in operational environments by measures which can more closely match a one second resolution or be administered without interruption.

Future Directions for Research

Throughout its progression, this program of research contributes a number of key implications to the practice of RPA training. Studies I and II addressed the questions of (1) is the HDD Menu Trainer able to increase student familiarity with the menus of the MQ-1 RPA, and (2) is the HDD Menu Trainer equally effective in blended or distance modalities? Results indicate the affirmative to both research questions and open further inquiry and avenues of study in blended learning. What aspects of the blended learning model are under examination today and, in turn, how might these efforts advance the quality and effectiveness of RPA training?

As blended, flipped, and hybrid learning models are increasingly expected within higher education curriculum (Saichaie, 2020), the future directions of inquiry specific to blended learning are focusing increasingly on learner engagement (Borup, Jensen, et al., 2020; Halverson & Graham, 2019). Halverson and Graham, (2019), for instance, situate their interests in emotional engagement relative to cognitive engagement by proposing that both, "…cognitive and emotional engagement are the key factors essential to understanding learner engagement" (p.

153). To this model of learner engagement, behavioral engagement would also be added, again in contrast with cognitive engagement;

> … attendance or participation in a class might be evidence of behavioral engagement, while indicators of focused attention or absorption would be evidence of cognitive engagement. Similarly, submitting a course assignment or spending time in a learning management system (LMS) would indicate behavioral engagement while evidence of a student's mental energy focused on asking questions, taking notes, checking understanding etc. would be evidence of cognitive engagement. (Borup, Graham, et al., 2020, p. 812)

Instruments to measure this conceptual framework of learner engagement and investigation of the mechanism which support high levels of student engagement are anticipated as blended, flipped, and hybrid learning models continue to be implemented. Research topics may include topics such as, (1) How do measures of engagement (i.e. cognitive, affective, and behavioral) predict academic success? (2) Does the effect of each engagement measure on academic success change across learning models? and/or (3) How do the engagement measures of Borup, Graham, et al. (2020) relate to one another? Understanding the role of engagement may better enable programs to successfully deliver learning outcomes within their classes and training courses.

Study III addressed the questions, (1) Will EEG-based metrics of cognitive states such as workload and task engagement be able to distinguish variation during the training of RPA pilots? and (2) Will ABM's cognitive metric for workload be associated with a self-reported workload metric like the NASA TLX? The results above support the former question using the B-Alert X-

24 wireless Bluetooth system of ABM and this new measure of performance could support several interests in fields such as augmented cognition and adaptive automation.

Part of the National Science Foundation, the Cognitive Neuroscience program is situated within the Division of Behavioral and Cognitive Sciences and, "… seeks to fund highly innovative proposals that employ brain-based measurements in order to advance our understanding of the neural systems that mediate cognitive processes." (National Science Foundation, 2021, para. 2) Areas of particular interest to the program include the cognitive processes of, attention, learning, memory, decision-making, language, social cognition, and emotions.

The cognitive process of learning is well aligned with the RPA training environment as well as the results of Study III. EEG collection at multiple points throughout training could assess whether the cognitive workload of a remote pilot changes over time when conducting the same procedure. In the same spirit of the procedures in Studies I and II, an assessment of learning or training effectiveness might include a mixed ANOVA measuring cognitive workload near the middle and the end of training, split according to autopilot assistance.

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + \varepsilon$$

Where;
Y is a measure of cognitive workload.
X_1 is lesson or collection time. $|_{\text{time} = 0,1}$
X_2 is autopilot assistance. $|_{\text{assistance} = 0,1}$
X_1X_2 is a possible interaction effect (time * automation)

The Cognitive Neuroscience program interest in attention may also support further investigation into the relationship between remote pilot cognitive workload and autopilot use in Study III. Aligned more closely to the interests of safety – than, perhaps, that of remote pilot training – a procedure which manipulates the availability of automation during simulated RPA operations could better document the effect of automation on workload and other cognitive states or processes.

Methodological and statistical advances are constantly increasing the potential of physiological signals like EEG (Belkhiria & Peysakhovich, 2020). In the efforts of Zhu et al. (2021), it is apparent that interest in developing and refining cognitive state metrics from smaller sets of EEG features remains an ongoing interest. Paired with these advances, technological advances – e.g. dry electrodes and in-ear EEG devices – will serve to expand access to EEG-based cognitive state metrics and encourage increasingly ecological environments for collection. Some have already leveraged these signals to predict human error (Baldwin et al., 2010; Mazaheri et al., 2009), while others have postulated how operator state monitoring may contribute to single-pilot or reduced-crew operations in aviation (Schmid & Stanton, 2020).

Studies I through III assessed adaptations in method (i.e. trainers and modalities) and measures of performance (e.g. workload and task engagement) in an effort to improve the efficiency and quality of remote pilot training in simulated MQ-1 operations. Even as the study of blended learning models begins to recognize cognitive engagement among its indicators of learner engagement (Borup, Graham, et al., 2020; Halverson & Graham, 2019), proponents of cognitive metrics – such as working memory capacity and attention – are proposing their application in intelligent instructional systems (Mathan & Yeung, 2015), or the real-time measurement of changes in cognitive workload in intelligent tutoring systems (ITS) (Mills et al., 2017). Fascinating work is on the horizon as the interests of instructional models and cognitive measurement increasingly intersect around the learning process.

References

Adler, M., & Johnson, K. (2000). *Quantifying the literature of computer-aided instruction in medical education*. Academic Medicine 75 (10).

Advanced Brain Monitoring. (2009). *B-alert Live: User Manual*.

Agius, R., & Bagnall, G. (1998). Development and evaluation of the use of the internet as an educational tool in occupational and environmental health and medicine. *Occupational Medicine*, *48*, 337–343.

Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., & Burrell, N. (2004). Evaluating the effectiveness of distance learning: a comparison using meta-analysis. *Journal of Communication*, *54*, 402–420.

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Pozzi, S., & Babiloni, F. (2016). A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks. *Progress in Brain Research*, *228*, 295–328.

Baldwin, C., Coyne, J. T., Roberts, D. M., Barrow, J. H., Cole, A., Sibley, C., Taylor, B., & Buzzell, G. (2010). *Prestimulus alpha as a precursor to errors in a UAV target orientation detection task.* Naval Research Lab.

Banks, E., Cook, S., Fredrick, G., Gill, S., Gray, J., Larue, T., Milton, J., Tootle, A., Wheeler, P., Snyder, P., & Waller, Z. (2018). NCHRP Project 20-68A, Scan 17-01: Successful Approaches for the Use of Unmanned Aerial System By Surface Transportation Agencies.

In *National Cooperative Highway Research Program*.

Baumlin, K., Bessette, M., Lewis, C., & Richardson, L. (2000). EMCyberSchool: An evaluation of computer-assisted instruction on the internet. *Academic Emergency Medicine*, *7*, 959–962.

Bedwell, W., & Salas, E. (2010). *Computer-based training: capitalizing on lessons learned*. International Journal of Training and Development 14 (3).

Belkhiria, C., & Peysakhovich, V. (2020). Electro-Encephalography and Electro-Oculography in Aeronautics: A Review Over the Last Decade (2010–2020). *Frontiers in Neuroergonomics*, *1*, 1–25. https://doi.org/doi: 10.3389/fnrgo.2020.606719

Bell, D., Fonarow, G., Hays, R., & Mangione, C. (2000). Self-study from web-based and printed guidline materials: A randomized controlled trial among resident physicians. *Annals of Internal Medicine*, *132*, 938–946.

Berka, C., J., L. D., N., L. M., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5), B231–B244.

Bernhardt, K. A., Poltavski, D., Petros, T., Ferraro, F. R., Jorgenson, T., Carlson, C., & Iseminger, C. (2019). The effects of dynamic workload and experience on commercially available EEG cognitive state metrics in a high-fidelity air traffic control environment. *Applied Ergonomics*, *77*, 83–91.

Block, A., Felix, M., Udermann, B., Reineke, D., & Murray, S. (2008). Achievement and satisfaction in an online versus a traditional health and wellness course. *Journal of Online Learning and Teaching*, *4*, 57–66.

Borghini, G., Arico`, P., Di Flumeri, G., Salinari, S., Colosimo, A., Bonelli, S., Napoletano, L., Ferreira, A., & Babiloni, F. (2015). Avionic technology testing by using a cognitive neurometric index: A study with professional helicopter pilots. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6182–6185. https://doi.org/http://dx.doi.org/10.1109/EMBC.2015.7319804

Borghini, G., Aricò, P., Ferri, F., Graziani, I., Pozzi, S., Napoletano, L., Imbert, J. P., Granger, G., Benhacene, R., & Babiloni, F. (2014). A neurophysiological training evaluation metric for air traffic management. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3005–3008. https://doi.org/http://dx.doi.org/10.1109/EMBC.2014.6944255

Borghini, G., Vecchiato, G., Toppi, J., Astolfi, L., Maglione, A., Isabella, R., Caltagirone, C., Kong, W., Wei, D., Zhou, Z., Polidori, L., Vitiello, S., & Babiloni, F. (2012). Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6442–6445. https://doi.org/http://dx.doi.org/10.1109/EMBC.2012.6347469

Borup, J., Graham, C. R., West, R. E., Archambault, L., & Spring, K. J. (2020). Academic Communities of Engagement: an expansive lens for examining support structures in blended and online learning. *Educational Technology Research and Development*, *68*(2), 807–832.

Borup, J., Jensen, M., Archambault, L., Short, C. R., & Graham, C. (2020). Supporting students during COVID-19: Developing and leveraging academic communities of engagement in a time of crisis. *Journal of Technology and Teacher Education*, *28*(2), 161–169.

Boyle, T., Bradley, C., Chalk, P., Jones, R., & Pickard, P. (2003). *Using blended learning to improve student success rates in learning to program*. Journal of Educational Media 28 (2&3).

Burke, M., Sarpy, S. A., Smith-Crowe, K., Chan-Serafin, S., Salvador, R., & Islam, G. (2006). Relative Effectiveness of Worker Safety and Health Training Methods. *American Journal of Public Health*, *96*, 315–324.

Carbonell, K. B., Dailey-Hebert, A., & Gihselaers. (2013). Unleashing the creative potential of faculty to create blended learning. *Internet and Higher Education*, *18*, 29–37.

Causse, M., Imbert, J. P., Giraudet, L., Jouffrais, C., & Tremblay, S. (2016). The role of cognitive and perceptual loads in inattentional deafness. *Frontiers in Human Neuroscience*, *10*, 1–12. https://doi.org/10.3389/fnhum.2016.00344

Chumley-Jones, H., Dobbie, A., & Alford, C. (2002). Web-based learning: sound educational method or hype? A review of the evaluative literature. *Academic Medicine*, *77*, 86–93.

Cohen, V. (2003). A model for assessing distance learning instruction. *Journal of Computing in Higher Education*, *14*, 98–120.

Curran, V., Hoekman, T., Gulliver, W., Landells, I., & Hatcher, L. (2000). Web-based continuing medical education (II): Evaluation study of computer-mediated continuing medical education. *The Journal of Continuing Education in the Health Profession*, *20*, 106–

119.

Davidson, R. J., Jackson, D. C., & Larson, C. L. (2000). Human electroencephalography. In *Handbook of Psychophysiology* (pp. 27–52).

DeBourgh, G. (1999). Technology is the tool teaching is the task: Student satisfaction in distance learning. *Society for Information Technology and Teacher Education International Conference*, 1–8.

DeRouin, R., Fritzsche, B., & Salas, E. (2004). Optimizing e-learning: reserach-based guidelines for learner-controlled training. *Human Resource Management Journal*, *43*, 147–162.

Engel, S., Crandall, J., Basch, C., Zybert, P., & Wylie-Rosett, J. (1997). Computer-assisted diabetes nutrition education increases knowledge and self-efficacy of medical students. *The Diabetes Educator*, *23*, 545–549.

Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage Publications.

Francis, B., Mauriello, S., Phillips, C., Englebardt, S., & Grayden, S. (2000). Assessment of online continuing dental education in north carolina. *The Journal of Continuing Education in the Health Profession*, *20*, 76–84.

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, *10*(9), 829–839.

Graham, C., Allen, S., & Ure, D. (2005). Benefits and Challenges of Blended Learning Environments. In *M* (pp. 253–259). IGI Global.

Graham, C., & Dziuban, C. (2008). Blended Learning Environments. In *J* (pp. 269–276).

Graham, C., Woodfield, W., & Harrison, B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *Internet and Higher Education*, *18*, 4–14.

Halverson, L. R., & Graham, C. (2019). Learner engagement in blended learning environments: A conceptual framework. *Online Learning*, *23*(2), 145–178.

Harris, J., Salasche, S., & Harris, R. (2001). Can internet-based continuing medical education improve physicians' skin cancer knowledge and skills? *Journal of General Internal Medicine*, *16*, 50–56.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183. https://doi.org/https://doi.org/10. 1016/S0166-4115(08)62386-9

Horsch, A., Balbach, T., Melnitzki, S., & Knauth, J. (2000). Learning tumor diagnostics and medical image processing via the WWW - The case-based radiological textbook ODITEB. *International Journal of Medical Informatics*, *58–59*, 39–50.

Howell, D. C. (2016). *Fundamental statistics for the behavioral sciences*. Cengage learning.

Hsu, L.-L., & Hsieh, S.-I. (2011). Effects of a blended learning module on self-reported learning performances in baccalaureate nursing students. *Journal of Advanced Nursing*, 2435–2444.

Irvine, V. (2020). The Landscape of Merging Modalities. *EDUCAUSE Review*, *55*(4).

Jenkins, D., & Vasigh, B. (2013). *The Economic Impact of Unmanned Aircraft Systems*

Integration in the United States. https://higherlogicdownload.s3.amazonaws.com/AUVSI/958c920a-7f9b-4ad2-9807-f9a4e95d1ef1/UploadedImages/New_Economic Report 2013 Full.pdf

John D. Odegard School of Aerospace Sciences. (2021). *Individual Course Training Costs*. https://aero.und.edu/aviation/student-info/individual-course-training-costs.html

Johnson, R. R., Popovic, D. P., Olmstead, R. E., Stikic, M., Levendowski, D. J., & Berka, C. (2011). Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological Psychology*, *87*(2), 241–250. https://doi.org/https://doi.org/10.1016/j.biopsycho.2011.03.003

Kronz, J., Silberman, M., Jr., A., W., & Epstein, J. (2000). A web-based tutorial improves practicing pathologists' gleason grading of images of prostate carcinoma specimens obtained by needle biopsy: Validation of a new medical education paradigm. *The American Cancer Society*, *89*(8), 1818–1823.

Laumakis, M., Graham, C., & Dziuban, C. (2009). The Sloan-C Pillars and Boundary Objects As a Framework for Evaluating Blended Learning. *Journal of Asynchronous Learning Networks*, *13*(1), 75–87.

Lipman, A., Sade, R., Glotzbach, A., Lancaster, C., & Marshall, M. (2001). The incremental value of internet-based instruction as an adjunct to classroom instruction: A prospective randomized study. *Academy of Medicine*, *76*(10), 72–76.

Lopez-Perez, M., Perez-Lopez, M., & Rodriguez-Ariza, L. (2011). Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers and Education*, *56*, 818–826.

Lutte, R. K., & Lovelace, K. (2016). Airline pilot supply in the US: Factors influencing the collegiate pilot pipeline. *Journal of Aviation Technology and Engineering*, *6*(1), 53–63.

Macdonald, J. S., Mathan, S., & Yeung, N. (2011). Trial-by-trial variations in subjective attentional state are reflected in ongoing prestimulus EEG alpha oscillations. *Frontiers in Psychology*, *2*, 1–16.

Mallat, C., Cegarra, J., Calmettes, C., & Capa, R. L. (2020). A curvilinear effect of mental workload on mental effort and behavioral adaptability: an approach with the pre-ejection period. *Human Factors*, *62*(6), 928–939. https://doi.org/10.1177/0018720819855919

Mandrick, K., Peysakhovich, V., Rémy, F., Lepron, E., & Causse, M. (2016). Neural and psychophysiological correlates of human performance under stress and high mental workload. *Biological Psychology*, *121*, 62–73.

Marcel, S., & Millán, J. D. R. (2007). Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 743–752. https://doi.org/http://dx.doi.org/10.1109/TPAMI.2007.1012

Mathan, S., & Yeung, N. (2015). *Extending the diagnostic capabilities of artificial intelligence-based instructional systems.* 51–60.

Mazaheri, A., Nieuwenhuis, I. L., Van Dijk, H., & Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Human Brain Mapping*, *30*(6), 1791–

1800.

Melton, B., Graf, H., & Chopak-Foss, J. (2009). Achievement and satisfaction in blended learning versus traditional general health course designs. *International Journal for the Scholarship of Teaching and Learning*, *3*, 1–13.

Miller, P. (2019). UAS degree program reaches 10-year anniversary. *UND Today*. http://blogs.und.edu/und-today/2019/12/uas-degree-program-reaches-10-year-milestone/

Mills, C., Fridman, I., Soussou, W., Waghray, D., Olney, A. M., & D'Mello, S. K. (2017). Put your thinking cap on: detecting cognitive load using EEG during learning. *Seventh International Learning Analytics & Knowledge Conference*, 80–89.

Mortera-Gutierrez, F. (2006). Faculty best practices using blended learning in e-learning and face-to-face instruction. *International Journal on E-Learning*, *5*, 313–337.

Moskal, P., Dziuban, C., & Hartman, J. (2013). Blended learning: A dangerous idea? *Internet and Higher Education*, *18*, 15–23.

National Science Foundation. (2021). *Cognitive Neuroscience (CogNeuro)*. https://beta.nsf.gov/funding/opportunities/cognitive-neuroscience-cogneuro

Norberg, A., Dziuban, C., & Moskal, P. (2011). A Time Based Blended Learning Model. *On the Horizon*, *19*(3), 207–216. https://doi.org/http://dx.doi.org/10.1108/10748121111163913

North Dakota Aeronautics Commisson. (2010). *North Dakota Economic Impact of Aviation*. http://library.nd.gov/statedocs/Aeronautics/EconomicImpactReport20120329.pdf

North Dakota Aeronautics Commisson. (2015). *Statewide Economic Impact of Aviation in North*

Dakota. https://aero.nd.gov/image/cache/Full_Technical_Report_-_FINAL_2.pdf

Osguthorpe, R., & Graham, C. (2003). Blended learning environments: Definitions and directions. *The Quarterly Review of Distance Education*, *4*(3), 227–233.

Overbaugh, R. (1994). *Reserach-based guidelines for computer-based instruction development*. Journal of Reserach on Computing in Education 27 (1).

Owston, R., York, D., & Murtha, S. (2013). Student perceptions and achievement in a university blended learning strategic initiative. *Internet and Higher Education*, *18*, 38–46.

Parasuraman, R. (2015). Neuroergonomic perspectives on human systems integration: Mental workload, vigilance, adaptive automation, and training. In D. A. Boehm-Davis, F. T. Durso, & J. D. Lee (Eds.), *APA handbook of human systems integration* (pp. 163–176). American Psychological Association.

Pereira, J., Pleguezuelos, E., Meri, A., Molina-Ros, A., Molina-Tomas, C., & Masdeu, C. (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical Education*, *41*, 189–195.

Perryer, G., Walmsley, A., Barclay, C., Shaw, L., & Smith, A. (2002). Development and evaluation of a stand-alone web-based CAL program. *European Journal of Dental Education*, *4*(3), 118–123.

Porter, W., Graham, C., Spring, K., & Welch, K. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers and Education*, *75*, 185–195.

Rivera, J., & Rice, M. (2002). A comparison of student outcomes & satisfaction between traditional & web based course offerings. *Online Journal of Distance Learning*

Administration, *5*.

Rose, M., Frisby, A., Hamlin, M., & Jones, S. (2000). Evaluation of the effectiveness of a web-based graduate epidemiology course. *Computers in Nursing*, *18*, 162–167.

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. http://www.rstudio.com/

Sakowski, H., Rich, E., & Turner, P. (2001). Web-based case simulations for a primary care clerkship. *Academic Medicine*, *76*(547.).

Schmid, D., & Stanton, N. A. (2020). Progressing Toward Airliners' Reduced-Crew Operations: A Systematic Literature Review. *The International Journal of Aerospace Psychology*, *30*(1–2), 1–24. https://doi.org/10.1080/24721840.2019.1696196

Schubert, R., Tangermann, M., Haufe, S., Sannelli, C., Simon, M., Schmidt, E. A., Kincsesc, W. E., & Curio, G. (2008). Parieto-occipital alpha power indexes distraction during simulated car driving. *International Journal of Psychophysiology*, *69*(3), 214. https://doi.org/http://dx.doi.org/10.1016/j.ijpsycho.2008.05.033

Sciarini, L. W., Grubb, J. D., & Fatolitis, P. G. (2014). Cognitive state assessment: Examination of EEG-based measures on a stroop task. *Human Factors and Ergonomics Society Annual Meeting*, 215–219.

Smyth, S., Houghton, C., Cooney, A., & Casey, D. (2012). Sutdents' experiences of blended learning across a range of postgraduate programmes. *Nurse Education Today*, *32*, 464–468.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604.

https://doi.org/https://doi.org/10.2466/ pms.1978.47.2.599

Venthur, B., Blankertz, B., Gugler, M. F., & Curio, G. (2010). Novel applications of BCI technology: psychophysiological optimization of working conditions in industry. *2010 IEEE International Conference on Systems, Man and Cybernetics*, 417–421. https://doi.org/10.1109/ICSMC.2010.5641772

Waller, Z., & Bridewell, J. (2014). Implementing UAS Programs: An Integrated Approach. *Collegiate Aviation News*, 12–13.

Waller, Z., Petros, T., Bridewell, J., Kroeber, S., & Nowatzki, N. (2016). Medium Altitude Long Endurance Remotely Piloted Aircraft Training: A Pilot Study in Blended Learning. *Journal of Unmanned Aerial Systems*, *1*, 6–16. https://doi.org/10.31356/avi-fac0002

Wechsler, D. (1997). WAIS III: Wechsler adult intelligence scale. *Psychological Corporation*.

Welke, S., Jurgensohn, T., & Roetting, M. (2009). Single-trial detection of cognitive processes for increasing traffic safety. *International Technical Conference on the Enhanced Safety of Vehicles*, 1–10.

Williams, K. (2004). *A Summary of Unmanned Aircraft Accident/Incident Data: Human Factors Implications*.

Williams, K. (2006). *Human Factors Implications of Unmanned Aircraft Accidents: Flight-Control Problems*.

Woo, M., & Kimmick, J. (2000). Comparison of internet versus lecture instructional methods for teaching nursing research. *Journal of Professional Nursing*, *16*, 132–139.

Yuan, Y., Chang, K. M., Taylor, J. N., & Mostow, J. (2014). Toward unobtrusive measurement of reading comprehension using low-cost EEG. *Fourth International Conference on Learning Analytics and Knowledge*, 54–58.

Zhu, G., Zong, F., Zhang, H., Wei, B., & Liu, F. (2021). Cognitive Load During Multitasking Can Be Accurately Assessed Based on Single Channel Electroencephalography Using Graph Methods. IEEE Access, 9, 33102–33109. https://doi.org/10.1109/ACCESS.2021.3058271

Appendix A
Codebooks for Studies I-III

Codebook for Study I

Demographic variables.

Name	Item
Age	Please state your age. [text box]
Gender	What is your gender? Circle an answer. (0) Male (1) Female
Education	What is your education level? Circle all that apply (0) High School (1) College Freshman (2) College Sophomore (3) College Junior (4) College Senior (5) Associate Degree (6) Bachelor Degree (7) Master’s Degree (8) Doctoral Degree (9) Other Post Graduate Advanced Certification
Ethnicity	What is your ethnic identification? (0) White (1) Black or African American (2) American Indian or Alaska Native (3) Asian (4) Native Hawaiian and Other Pacific Islander (5) Other

Aviation experience variables.

Name	Item
Pilot	Do you hold an FAA Pilot Certificate? (0) No (1) Yes
Certificate	If so, what Pilot Certificate do you hold? (0) None (1) Private (2) Commercial (3) Airline Transport Pilot

Hours_Total	If so, what is your total flight time? (0) 0-50 hours (1) 50-100 hours (2) 100-150 hours (3) 150-200 hours (4) 200-300 hours (5) 300-500 hours (6) 500-1000 hours (7) More than 1000 hours
Medical	Do you hold an FAA Aviation Medical Certificate? (0) No (1) Yes
Class	If so, what class is your Aviation Medical Certificate? (1) First (2) Second (3) Third (4) Expired
Instrument	Do you hold an instrument rating? (0) No (1) Yes
Hours_Inst_ Actual	Please estimate your total instrument time [text box]
Hours_Inst_ Simulated	Please estimate your total number of simulated instrument time (i.e. hood time) [text box]
Hours_Inst_ FTD	Please estimate your total number of hours in a Ground Training Device or Aircraft Simulator. [text box]

Pedagogy and learning efficiency variables.

Name	Item
Pedagogy	Pedagogy used to deliver HDD instruction. (0) Traditional (1) Distance (2) Blended
Study	Hours of self-study reported by the participant. [text box]
Percent_Gain	Percent Change from Pre-test to Post-test Score

Learner knowledge variables – HDD menu trainer pretest.

Name	Item
Pre_Score	Participant's Pretest Score [text box]
Pre_Inc_1	Number of incorrect key strokes on HDD task 1 [text box]
Pre_Inc_2	Number of incorrect key strokes on HDD task 2 [text box]
Pre_Inc_3	Number of incorrect key strokes on HDD task 3 [text box]

. . .	
Pre_Inc_25	Number of incorrect key strokes on HDD task 25 [text box]
Pre_Skip_1	HDD task 1 was skipped. (0) No (1) Yes
Pre_Skip_2	HDD task 2 was skipped. (0) No (1) Yes
Pre_Skip_3	HDD task 3 was skipped. (0) No (1) Yes
. . .	
Pre_Skip_25	HDD task 25 was skipped. (0) No (1) Yes
Pre_Time_1	Elapsed time – in seconds – on HDD task 1 [text box]
Pre_Time_2	Elapsed time – in seconds – on HDD task 2 [text box]
Pre_Time_3	Elapsed time – in seconds – on HDD task 3 [text box]
. . .	
Pre_Time_25	Elapsed time – in seconds – on HDD task 25 [text box]

Learner knowledge variables – HDD menu trainer posttest.

Name	**Item**
Post_Score	Participant's Posttest Score [text box]
Post_Inc_1	Number of incorrect key strokes on HDD task 1 [text box]
Post_Inc_2	Number of incorrect key strokes on HDD task 2 [text box]
Post_Inc_3	Number of incorrect key strokes on HDD task 3 [text box]
. . .	
Post_Inc_25	Number of incorrect key strokes on HDD task 25 [text box]

Post_Skip_1	HDD task 1 was skipped. (0) No (1) Yes
Post_Skip_2	HDD task 2 was skipped. (0) No (1) Yes
Post_Skip_3	HDD task 3 was skipped. (0) No (1) Yes
. . .	
Post_Skip_25	HDD task 25 was skipped. (0) No (1) Yes
Post_Time_1	Elapsed time – in seconds – on HDD task 1 [text box]
Post_Time_2	Elapsed time – in seconds – on HDD task 2 [text box]
Post_Time_3	Elapsed time – in seconds – on HDD task 3 [text box]
. . .	
Post_Time_25	Elapsed time – in seconds – on HDD task 25 [text box]

Codebook for Study II

Demographic variables.

Name	Item
Age	Please state your age. [text box]
Gender	What is your gender? Circle an answer. (0) Male (1) Female
Education	What is your education level? Circle all that apply (0) High School (1) College Freshman (2) College Sophomore (3) College Junior (4) College Senior (5) Associate Degree (6) Bachelor Degree (7) Master's Degree (8) Doctoral Degree (9) Other Post Graduate Advanced Certification
Ethnicity	What is your ethnic identification? (0) White (1) Black or African American (2) American Indian or Alaska Native (3) Asian (4) Native Hawaiian and Other Pacific Islander (5) Other

Aviation experience variables.

Name	Item
Pilot	Do you hold an FAA Pilot Certificate? (0) No (1) Yes
Certificate	If so, what Pilot Certificate do you hold? (0) None (1) Private (2) Commercial (3) Airline Transport Pilot
Hours_Total	If so, what is your total flight time? (0) 0-50 hours (1) 50-100 hours (2) 100-150 hours (3) 150-200 hours (4) 200-300 hours (5) 300-500 hours (6) 500-1000 hours (7) More than 1000 hours

Medical	Do you hold an FAA Aviation Medical Certificate? (0) No (1) Yes
Class	If so, what class is your Aviation Medical Certificate? (1) First (2) Second (3) Third (4) Expired
Instrument	Do you hold an instrument rating? (0) No (1) Yes
Hours_Inst_ Actual	Please estimate your total instrument time [text box]
Hours_Inst_ Simulated	Please estimate your total number of simulated instrument time (i.e. hood time) [text box]
Hours_Inst_ FTD	Please estimate your total number of hours in a Ground Training Device or Aircraft Simulator. [text box]

Modality and learning efficiency variables.

Name	**Item**
Modality	Modality used to deliver HDD instruction. (0) Traditional (1) Distance (2) Blended
Study	Hours of self-study reported by the participant. [text box]
Percent_Gain	Percent Change from Pre-test to Post-test Score

Learner knowledge variables – HDD menu trainer pretest.

Name	**Item**
Pre_Score	Participant's Pretest Score [text box]
Pre_Inc_1	Number of incorrect key strokes on HDD task 1 [text box]
Pre_Inc_2	Number of incorrect key strokes on HDD task 2 [text box]
Pre_Inc_3	Number of incorrect key strokes on HDD task 3 [text box]
. . .	
Pre_Inc_25	Number of incorrect key strokes on HDD task 25 [text box]
Pre_Skip_1	HDD task 1 was skipped. (0) No (1) Yes

Pre_Skip_2	HDD task 2 was skipped. (0) No (1) Yes
Pre_Skip_3	HDD task 3 was skipped. (0) No (1) Yes
. . .	
Pre_Skip_25	HDD task 25 was skipped. (0) No (1) Yes
Pre_Time_1	Elapsed time – in seconds – on HDD task 1 [text box]
Pre_Time_2	Elapsed time – in seconds – on HDD task 2 [text box]
Pre_Time_3	Elapsed time – in seconds – on HDD task 3 [text box]
. . .	
Pre_Time_25	Elapsed time – in seconds – on HDD task 25 [text box]

Learner knowledge variables – HDD menu trainer posttest.

Name	Item
Post_Score	Participant's Posttest Score [text box]
Post_Inc_1	Number of incorrect key strokes on HDD task 1 [text box]
Post_Inc_2	Number of incorrect key strokes on HDD task 2 [text box]
Post_Inc_3	Number of incorrect key strokes on HDD task 3 [text box]
. . .	
Post_Inc_25	Number of incorrect key strokes on HDD task 25 [text box]
Post_Skip_1	HDD task 1 was skipped. (1) No (2) Yes
Post_Skip_2	HDD task 2 was skipped. (0) No (1) Yes
Post_Skip_3	HDD task 3 was skipped. (0) No (1) Yes

. . .	
Post_Skip_25	HDD task 25 was skipped. (0) No (1) Yes
Post_Time_1	Elapsed time – in seconds – on HDD task 1 [text box]
Post_Time_2	Elapsed time – in seconds – on HDD task 2 [text box]
Post_Time_3	Elapsed time – in seconds – on HDD task 3 [text box]
. . .	
Post_Time_25	Elapsed time – in seconds – on HDD task 25 [text box]

Learner attitude variables.

"Below are a number of statements, which may or may not apply to you regarding the coursework you have recently completed. For example, I felt well prepared for the final assessment*? Please write a number next to each statement which indicates the extent to which you agree or disagree with that statement."*

1 (Strongly disagree) to 5 (Strongly agree)

Name	**Item**
Att_Sum	Sum of participant's responses regarding the course.
Att_1	I felt I understood the subject well.
Att_2	Course material was presented in an appropriate and effective way.
Att_3	Presentation of course material kept my attention.
Att_4	I was motivated to work and learn in this course.
Att_5	I was satisfied with the pace that material was presented to me.
Att_6	I was satisfied with the amount and availability of instructor feedback.
Att_7	I gained a satisfactory amount of knowledge regarding the course topic.
Att_8	I felt well prepared for the final assessment.
	Please describe specific aspects of the course or instruction which promoted your learning. [text box]
	Please describe improvements, if any, which would better assist your learning of the course material. [text box]

Codebook for Study III

Demographic variables.

Name	Item
Age	Please state your age. [text box]
Gender	What is your gender? Circle an answer. (0) Male (1) Female
Education	What is your education level? Circle all that apply (0) High School (1) College Freshman (2) College Sophomore (3) College Junior (4) College Senior (5) Associate Degree (6) Bachelor Degree (7) Master's Degree (8) Doctoral Degree (9) Other Post Graduate Advanced Certification
Ethnicity	What is your ethnic identification? (0) White (1) Black or African American (2) American Indian or Alaska Native (3) Asian (4) Native Hawaiian and Other Pacific Islander (5) Other

Aviation experience variables.

Name	Item
Pilot	Do you hold an FAA Pilot Certificate? (0) No (1) Yes
Certificate	If so, what Pilot Certificate do you hold? (0) None (1) Private (2) Commercial (3) Airline Transport Pilot
Hours_Total	If so, what is your total flight time? (0) 0-50 hours (1) 50-100 hours (2) 100-150 hours (3) 150-200 hours (4) 200-300 hours (5) 300-500 hours (6) 500-1000 hours (7) More than 1000 hours

Medical	Do you hold an FAA Aviation Medical Certificate? (0) No (1) Yes
Class	If so, what class is your Aviation Medical Certificate? (1) First (2) Second (3) Third (4) Expired
Instrument	Do you hold an instrument rating? (0) No (1) Yes
Hours_Inst_ Actual	Please estimate your total instrument time [text box]
Hours_Inst_ Simulated	Please estimate your total number of simulated instrument time (i.e. hood time) [text box]
Hours_Inst_ FTD	Please estimate your total number of hours in a Ground Training Device or Aircraft Simulator. [text box]

Mental rotation and WAIS III

Name	Item
MRT	Participant's Score on the Mental Rotation A or B [text box]
WAIS_III	Participant's Score on the Wechsler Adult Intelligence Scale-Third Edition [text box]

Learner knowledge variables – HDD menu trainer pretest.

Name	Item
Pre_Score	Participant's Pretest Score [text box]
Pre_Inc_1	Number of incorrect key strokes on HDD task 1 [text box]
Pre_Inc_2	Number of incorrect key strokes on HDD task 2 [text box]
Pre_Inc_3	Number of incorrect key strokes on HDD task 3 [text box]
. . .	
Pre_Inc_25	Number of incorrect key strokes on HDD task 25 [text box]
Pre_Skip_1	HDD task 1 was skipped. (0) No (1) Yes
Pre_Skip_2	HDD task 2 was skipped. (0) No (1) Yes

Pre_Skip_3	HDD task 3 was skipped. (0) No (1) Yes
. . .	
Pre_Skip_25	HDD task 25 was skipped. (0) No (1) Yes
Pre_Time_1	Elapsed time – in seconds – on HDD task 1 [text box]
Pre_Time_2	Elapsed time – in seconds – on HDD task 2 [text box]
Pre_Time_3	Elapsed time – in seconds – on HDD task 3 [text box]
. . .	
Pre_Time_25	Elapsed time – in seconds – on HDD task 25 [text box]

Learner knowledge variables – HDD menu trainer posttest.

Name	**Item**
Post_Score	Participant's Posttest Score [text box]
Post_Inc_1	Number of incorrect key strokes on HDD task 1 [text box]
Post_Inc_2	Number of incorrect key strokes on HDD task 2 [text box]
Post_Inc_3	Number of incorrect key strokes on HDD task 3 [text box]
. . .	
Post_Inc_25	Number of incorrect key strokes on HDD task 25 [text box]
Post_Skip_1	HDD task 1 was skipped. (0) No (3) Yes
Post_Skip_2	HDD task 2 was skipped. (0) No (1) Yes
Post_Skip_3	HDD task 3 was skipped. (0) No (1) Yes
. . .	

Post_Skip_25	HDD task 25 was skipped. (0) No (1) Yes
Post_Time_1	Elapsed time – in seconds – on HDD task 1 [text box]
Post_Time_2	Elapsed time – in seconds – on HDD task 2 [text box]
Post_Time_3	Elapsed time – in seconds – on HDD task 3 [text box]
. . .	
Post_Time_25	Elapsed time – in seconds – on HDD task 25 [text box]

NASA task load index.

Name	**Item**
TLX_T1_PW_Mental	Sum of Participant's pairwise selection for Mental Demand at time 1. [text box – range 1:5]
TLX_T1_PW_Physical	Sum of Participant's pairwise selection for Physical Demand at time 1. [text box – range 1:5]
TLX_T1_PW_Temporal	Sum of Participant's pairwise selection for Temporal Demand at time 1. [text box – range 1:5]
TLX_T1_PW_Performance	Sum of Participant's pairwise selection for Performance at time 1. [text box – range 1:5]
TLX_T1_PW_Effort	Sum of Participant's pairwise selection for Effort at time 1. [text box – range 1:5]
TLX_T1_PW_Frustration	Sum of Participant's pairwise selection for Frustration at time 1. [text box – range 1:5]
TLX_T1_RS_CL_Mental	Participant's raw score for Mental Demand while completing the Checklist at Time 1. [text box – range 0:100]
TLX_T1_RS_CL_Physical	Participant's raw score for Physical Demand while completing the Checklist at Time 1. [text box – range 0:100]
TLX_T1_RS_CL_Temporal	Participant's raw score for Temporal Demand while completing the Checklist at Time 1. [text box – range 0:100]
TLX_T1_RS_CL_Performance	Participant's raw score for Performance while completing the Checklist at Time 1. [text box – range 0:100]
TLX_T1_RS_CL_Effort	Participant's raw score for Effort while completing the Checklist at Time 1. [text box – range 0:100]

TLX_T1_RS_CL_Frustration	Participant's raw score for Frustration while completing the Checklist at Time 1. [text box – range 0:100]
TLX_T1_AR_CL_Mental	Participant's adjusted rating for Mental Demand while completing the Checklist at Time 1. [text box – range 0:500]
TLX_T1_AR_CL_Physical	Participant's adjusted rating for Physical Demand while completing the Checklist at Time 1. [text box – range 0:500]
TLX_T1_AR_CL_Temporal	Participant's adjusted rating for Temporal Demand while completing the Checklist at Time 1. [text box – range 0:500]
TLX_T1_AR_CL_Performance	Participant's adjusted rating for Performance while completing the Checklist at Time 1. [text box – range 0:500]
TLX_T1_AR_CL_Effort	Participant's adjusted rating for Effort while completing the Checklist at Time 1. [text box – range 0:500]
TLX_T1_AR_CL_Frustration	Participant's adjusted rating for Frustration while completing the Checklist at Time 1. [text box – range 0:500]
TLX_T1_WR_CL	Participant's weighted rating for workload while completing the Checklist at Time 1. [text box]
TLX_T1_RS_FP_Mental	Participant's raw score for Mental Demand while completing the Flight Pattern at Time 1. [text box – range 0:100]
TLX_T1_RS_FP_Physical	Participant's raw score for Physical Demand while completing the Flight Pattern at Time 1. [text box – range 0:100]
TLX_T1_RS_FP_Temporal	Participant's raw score for Temporal Demand while completing the Flight Pattern at Time 1. [text box – range 0:100]
TLX_T1_RS_FP_Performance	Participant's raw score for Performance while completing the Flight Pattern at Time 1. [text box – range 0:100]
TLX_T1_RS_FP_Effort	Participant's raw score for Effort while completing the Flight Pattern at Time 1. [text box – range 0:100]
TLX_T1_RS_FP_Frustration	Participant's raw score for Frustration while completing the Flight Pattern at Time 1. [text box – range 0:100]
TLX_T1_AR_FP_Mental	Participant's adjusted rating for Mental Demand while completing the Flight Pattern at Time 1. [text box – range 0:500]
TLX_T1_AR_FP_Physical	Participant's adjusted rating for Physical Demand while completing the Flight Pattern at Time 1. [text box – range 0:500]

TLX_T1_AR_FP_Temporal	Participant's adjusted rating for Temporal Demand while completing the Flight Pattern at Time 1. [text box – range 0:500]
TLX_T1_AR_FP_Performance	Participant's adjusted rating for Performance while completing the Flight Pattern at Time 1. [text box – range 0:500]
TLX_T1_AR_FP_Effort	Participant's adjusted rating for Effort while completing the Flight Pattern at Time 1. [text box – range 0:500]
TLX_T1_AR_FP_Frustration	Participant's adjusted rating for Frustration while completing the Flight Pattern at Time 1. [text box – range 0:500]
TLX_T1_WR_FP	Participant's weighted rating for workload while completing the Flight Pattern at Time 1. [text box]
TLX_T2_PW_Mental	Sum of Participant's pairwise selection for Mental Demand at time 2. [text box – range 1:5]
TLX_T2_PW_Physical	Sum of Participant's pairwise selection for Physical Demand at time 2. [text box – range 1:5]
TLX_T2_PW_Temporal	Sum of Participant's pairwise selection for Temporal Demand at time 2. [text box – range 1:5]
TLX_T2_PW_Performance	Sum of Participant's pairwise selection for Performance at time 2. [text box – range 1:5]
TLX_T2_PW_Effort	Sum of Participant's pairwise selection for Effort at time 2. [text box – range 1:5]
TLX_T2_PW_Frustration	Sum of Participant's pairwise selection for Frustration at time 2. [text box – range 1:5]
TLX_T2_RS_CL_Mental	Participant's raw score for Mental Demand while completing the Checklist at Time 2. [text box – range 0:100]
TLX_T2_RS_CL_Physical	Participant's raw score for Physical Demand while completing the Checklist at Time 2. [text box – range 0:100]
TLX_T2_RS_CL_Temporal	Participant's raw score for Temporal Demand while completing the Checklist at Time 1. [text box – range 0:100]
TLX_T2_RS_CL_Performance	Participant's raw score for Performance while completing the Checklist at Time 2. [text box – range 0:100]
TLX_T2_RS_CL_Effort	Participant's raw score for Effort while completing the Checklist at Time 2. [text box – range 0:100]
TLX_T2_RS_CL_Frustration	Participant's raw score for Frustration while completing the Checklist at Time 2. [text box – range 0:100]

TLX_T2_AR_CL_Mental	Participant’s adjusted rating for Mental Demand while completing the Checklist at Time 2. [text box – range 0:500]
TLX_T2_AR_CL_Physical	Participant’s adjusted rating for Physical Demand while completing the Checklist at Time 2. [text box – range 0:500]
TLX_T2_AR_CL_Temporal	Participant’s adjusted rating for Temporal Demand while completing the Checklist at Time 2. [text box – range 0:500]
TLX_T2_AR_CL_Performance	Participant’s adjusted rating for Performance while completing the Checklist at Time 2. [text box – range 0:500]
TLX_T2_AR_CL_Effort	Participant’s adjusted rating for Effort while completing the Checklist at Time 2. [text box – range 0:500]
TLX_T2_AR_CL_Frustration	Participant’s adjusted rating for Frustration while completing the Checklist at Time 2. [text box – range 0:500]
TLX_T2_WR_CL	Participant’s weighted rating for workload while completing the Checklist at Time 2. [text box]
TLX_T2_RS_FP_Mental	Participant’s raw score for Mental Demand while completing the Flight Pattern at Time 2. [text box – range 0:100]
TLX_T2_RS_FP_Physical	Participant’s raw score for Physical Demand while completing the Flight Pattern at Time 2. [text box – range 0:100]
TLX_T2_RS_FP_Temporal	Participant’s raw score for Temporal Demand while completing the Flight Pattern at Time 2. [text box – range 0:100]
TLX_T2_RS_FP_Performance	Participant’s raw score for Performance while completing the Flight Pattern at Time 2. [text box – range 0:100]
TLX_T2_RS_FP_Effort	Participant’s raw score for Effort while completing the Flight Pattern at Time 2. [text box – range 0:100]
TLX_T2_RS_FP_Frustration	Participant’s raw score for Frustration while completing the Flight Pattern at Time 2. [text box – range 0:100]
TLX_T2_AR_FP_Mental	Participant’s adjusted rating for Mental Demand while completing the Flight Pattern at Time 2. [text box – range 0:500]
TLX_T2_AR_FP_Physical	Participant’s adjusted rating for Physical Demand while completing the Flight Pattern at Time 2. [text box – range 0:500]
TLX_T2_AR_FP_Temporal	Participant’s adjusted rating for Temporal Demand while completing the Flight Pattern at Time 2. [text box – range 0:500]
TLX_T2_AR_FP_Performance	Participant’s adjusted rating for Performance while completing the Flight Pattern at Time 2.

	[text box – range 0:500]
TLX_T2_AR_FP_Effort	Participant’s adjusted rating for Effort while completing the Flight Pattern at Time 2. [text box – range 0:500]
TLX_T2_AR_FP_Frustration	Participant’s adjusted rating for Frustration while completing the Flight Pattern at Time 2. [text box – range 0:500]
TLX_T2_WR_FP	Participant’s weighted rating for workload while completing the Flight Pattern at Time 2. [text box]

ABM cognitive state metrics – workload.

Name	Item
EEG_T1_CL_Workload.1	ABM’s probability of average workload in epoch 1 during the Checklist at Time 1. [text box]
EEG_T1_CL_Workload.2	ABM’s probability of average workload in epoch 2 during the Checklist at Time 1. [text box]
EEG_T1_CL_Workload.3	ABM’s probability of average workload in epoch 3 during the Checklist at Time 1. [text box]
. . .	
EEG_T1_CL_Workload.N	ABM’s probability of average workload in epoch N during the Checklist at Time 1. [text box]
EEG_T1_FP.1_Workload.1	ABM’s probability of average workload in epoch 1 during leg 1 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.1_Workload.2	ABM’s probability of average workload in epoch 2 during leg 1 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.1_Workload.3	ABM’s probability of average workload in epoch 3 during leg 1 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.1_Workload.N	ABM’s probability of average workload in epoch N during leg 1 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.2_Workload.1	ABM’s probability of average workload in epoch 1 during leg 2 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.2_Workload.2	ABM’s probability of average workload in epoch 2 during leg 2 of the Flight Pattern at Time 1. [text box]

EEG_T1_FP.2_Workload.3	ABM's probability of average workload in epoch 3 during leg 2 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.2_Workload.N	ABM's probability of average workload in epoch N during leg 2 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.3_Workload.1	ABM's probability of average workload in epoch 1 during leg 3 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.3_Workload.2	ABM's probability of average workload in epoch 2 during leg 3 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.3_Workload.3	ABM's probability of average workload in epoch 3 during leg 3 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.3_Workload.N	ABM's probability of average workload in epoch N during leg 3 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.4_Workload.1	ABM's probability of average workload in epoch 1 during leg 4 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.4_Workload.2	ABM's probability of average workload in epoch 2 during leg 4 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.4_Workload.3	ABM's probability of average workload in epoch 3 during leg 4 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.4_Workload.N	ABM's probability of average workload in epoch N during leg 4 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.5_Workload.1	ABM's probability of average workload in epoch 1 during leg 5 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.5_Workload.2	ABM's probability of average workload in epoch 2 during leg 5 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.5_Workload.3	ABM's probability of average workload in epoch 3 during leg 5 of the Flight Pattern at Time 1. [text box]

. . .	
EEG_T1_FP.5_Workload.N	ABM’s probability of average workload in epoch N during leg 5 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.6_Workload.1	ABM’s probability of average workload in epoch 1 during leg 6 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.6_Workload.2	ABM’s probability of average workload in epoch 2 during leg 6 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.6_Workload.3	ABM’s probability of average workload in epoch 3 during leg 6 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.6_Workload.N	ABM’s probability of average workload in epoch N during leg 6 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.7_Workload.1	ABM’s probability of average workload in epoch 1 during leg 7 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.7_Workload.2	ABM’s probability of average workload in epoch 2 during leg 7 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.7_Workload.3	ABM’s probability of average workload in epoch 3 during leg 7 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.7_Workload.N	ABM’s probability of average workload in epoch N during leg 7 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.8_Workload.1	ABM’s probability of average workload in epoch 1 during leg 8 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.8_Workload.2	ABM’s probability of average workload in epoch 2 during leg 8 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.8_Workload.3	ABM’s probability of average workload in epoch 3 during leg 8 of the Flight Pattern at Time 1. [text box]
. . .	

EEG_T1_FP.8_Workload.N	ABM’s probability of average workload in epoch N during leg 8 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.9_Workload.1	ABM’s probability of average workload in epoch 1 during leg 9 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.9_Workload.2	ABM’s probability of average workload in epoch 2 during leg 9 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.9_Workload.3	ABM’s probability of average workload in epoch 3 during leg 9 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.9_Workload.N	ABM’s probability of average workload in epoch N during leg 9 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.10_Workload.1	ABM’s probability of average workload in epoch 1 during leg 10 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.10_Workload.2	ABM’s probability of average workload in epoch 2 during leg 10 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.10_Workload.3	ABM’s probability of average workload in epoch 3 during leg 10 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.10_Workload.N	ABM’s probability of average workload in epoch N during leg 10 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.11_Workload.1	ABM’s probability of average workload in epoch 1 during leg 11 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.11_Workload.2	ABM’s probability of average workload in epoch 2 during leg 11 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.11_Workload.3	ABM’s probability of average workload in epoch 3 during leg 11 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.11_Workload.N	ABM’s probability of average workload in epoch N during leg 11 of the Flight Pattern at Time 1. [text box]

EEG_T1_FP.12_Workload.1	ABM's probability of average workload in epoch 1 during leg 12 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.12_Workload.2	ABM's probability of average workload in epoch 2 during leg 12 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.12_Workload.3	ABM's probability of average workload in epoch 3 during leg 12 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.12_Workload.N	ABM's probability of average workload in epoch N during leg 12 of the Flight Pattern at Time 1. [text box]

ABM cognitive state metrics – engagement.

Name	Item
EEG_T1_CL_Engagement.1	ABM's probability of high engagement in epoch 1 during the Checklist at Time 1. [text box]
EEG_T1_CL_Engagement.2	ABM's probability of high engagement in epoch 2 during the Checklist at Time 1. [text box]
EEG_T1_CL_Engagement.3	ABM's probability of high engagement in epoch 3 during the Checklist at Time 1. [text box]
. . .	
EEG_T1_CL_Engagement.N	ABM's probability of high engagement in epoch N during the Checklist at Time 1. [text box]
EEG_T1_FP.1_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 1 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.1_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 1 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.1_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 1 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.1_Engagement.N	ABM's probability of high engagement in epoch N during leg 1 of the Flight Pattern at Time 1. [text box]

EEG_T1_FP.2_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 2 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.2_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 2 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.2_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 2 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.2_Engagement.N	ABM's probability of high engagement in epoch N during leg 2 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.3_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 3 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.3_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 3 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.3_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 3 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.3_Engagement.N	ABM's probability of high engagement in epoch N during leg 3 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.4_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 4 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.4_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 4 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.4_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 4 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.4_Engagement.N	ABM's probability of high engagement in epoch N during leg 4 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.5_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 5 of the Flight Pattern at Time 1. [text box]

EEG_T1_FP.5_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 5 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.5_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 5 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.5_Engagement.N	ABM's probability of high engagement in epoch N during leg 5 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.6_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 6 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.6_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 6 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.6_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 6 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.6_Engagement.N	ABM's probability of high engagement in epoch N during leg 6 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.7_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 7 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.1_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 7 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.7_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 7 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.7_Engagement.N	ABM's probability of high engagement in epoch N during leg 7 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.8_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 8 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.8_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 8 of the Flight Pattern at Time 1. [text box]

EEG_T1_FP.8_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 8 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.8_Engagement.N	ABM's probability of high engagement in epoch N during leg 8 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.9_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 9 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.9_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 9 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.9_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 9 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.9_Engagement.N	ABM's probability of high engagement in epoch N during leg 9 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.10_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 10 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.10_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 10 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.10_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 10 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.10_Engagement.N	ABM's probability of high engagement in epoch N during leg 10 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.11_Engagement.1	ABM's probability of high engagement in epoch 1 during leg 11 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.11_Engagement.2	ABM's probability of high engagement in epoch 2 during leg 11 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.11_Engagement.3	ABM's probability of high engagement in epoch 3 during leg 11 of the Flight Pattern at Time 1. [text box]

. . .	
EEG_T1_FP.11_Engagement.N	ABM’s probability of high engagement in epoch N during leg 11 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.12_Engagement.1	ABM’s probability of high engagement in epoch 1 during leg 12 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.12_Engagement.2	ABM’s probability of high engagement in epoch 2 during leg 12 of the Flight Pattern at Time 1. [text box]
EEG_T1_FP.12_Engagement.3	ABM’s probability of high engagement in epoch 3 during leg 12 of the Flight Pattern at Time 1. [text box]
. . .	
EEG_T1_FP.12_Engagement.N	ABM’s probability of high engagement in epoch N during leg 12 of the Flight Pattern at Time 1. [text box]