AFRL FIGHT'S ON!

THE AIR FORCE RESEARCH LABORATORY

ISSUE 63
FALL 2021

SIMULATING THE FUTURE

The Warfighter of Cognition and Efficiency

I/ITSEC Edition

"THE EFFECTIVENESS OF AIR FORCE AIRPOWER COMES DIRECTLY FROM THE POWER OF AIRMEN. WHILE IT IS NATURAL TO DEFINE THE AIR FORCE IN TERMS OF ITS AIRCRAFT, MISSILES, OR SATELLITES, IN REALITY, THE SERVICE'S UNMATCHED CAPABILITIES EXIST ONLY AND PRECISELY BECAUSE OF THE IMAGINATION, INNOVATION, AND DEDICATION OF ITS PEOPLE."

— Global Vigilance, Global Reach, Global Power for America
(USAF Strategic Document)

AFRL
THE AIR FORCE RESEARCH LABORATORY
LEAD | DISCOVER | DEVELOP | DELIVER

Graphic Designer **Ms. Shania Horner, RHW**
Technical Editor **Mr. Sam Angelo, RHW**
Publication Coordinator **Mr. Will Graver, RHW**

Photo by Spc. Jessica Scott

TABLE OF CONTENTS

Photo by J.M. Eddins Jr.

RHW

WARFIGHTER INTERACTIONS AND READINESS DIVISION

The United States Air Force stands at the forefront of progress, regarding both state intelligence and the subsequent initiative that reveals the promising tomorrow that can—and will—be earned. Focusing on improving both preparation and operations needed for successful military operations, the Air Force Research Laboratory (AFRL) procures results through the Warfighter Interactions & Readiness Division (RHW). The Division is committed to establishing and securing the pathway to the Warfighter's legacy of success and future efficiency by working with other Divisions and various researchers to produce findings and subsequent training that improve the cognition and mental fortitude of our Airmen and Guardians. Through the combined efforts of the RHW, the USAF will continue to grow into a single, united force that unwaveringly pursues and overcomes any and every challenge confronted: *one team, one fight*.

U.S. AIR FORCE

AIR FORCE RESEARCH LABORATORY

711th HUMAN PERFORMANCE WING

AIRMAN SYSTEMS DIRECTORATE

RHW
Technical Organization

WARFIGHTER INTERACTIONS & READINESS DIVISION 711 HPW/RHW

ACTING DIVISION CHIEF
Col Alfredo Rivera

WARFIGHTER INTERACTIONS & READINESS DIVISION
711 HPW/RHW

CORE TECHNICAL COMPETENCIES

TRAINING CTC
Dr. Glenn Gunzelmann

COGNITIVE MODELS CRA
Dr. Megan Morris

PERSONALIZED LEARNING & READINESS SCIENCES CRA
Ms. Jennifer Winner

ADAPTIVE WARFIGHTER INTERFACES CTC
Dr. Mark Draper

MULTISENSORY PERCEPTION & COMMUNICATION CRA
Dr. Griffin Romigh

SYSTEMS ANALYTICS CRA
Dr. Vincent Schmidt

COLLABORATIVE INTERFACES & TEAMING CRA
Dr. Chris Brill

PRODUCT LINES

READINESS PL
Dr. Winston "Wink" Bennett

ANALYTIC TOOLS PL
Mr. Greg Beister

AIRMAN-MACHINE INTEGRATION PL
Dr. Deirdre Mahle

RHW
Branch Structure

RHWA

MISSION ANALYTICS
Dr. William Murdock

RHWC

COLLABORATIVE INTERFACES AND TEAMING
Dr. Timothy Webb

RHWL

CONTINUOUS LEARNING
Lt Col James Lievsay

RHWM

COGNITIVE MODELS
Lt Col Christopher Terpening

RHWS

SENSORY SYSTEMS
Ms. Jennifer Brown

RHW

WARFIGHTER INTERACTIONS AND READINESS DIVISION

COL ALFREDO RIVERA
Acting Division Chief,
711 HPW/RHW

Welcome to the 2021 issue of fall *Fight's ON!* It is my honor to work in the company of the talented and professional members of the Warfighter Interactions & Readiness Division, RHW. I'm the acting Division Chief, Alfredo Rivera, an RC-135 Electronic Warfare Officer with experience in the Air Operations Center and CNAF (Commander, Naval Air Force) staff for operations, exercises, plans, programs, and International Relations. I'm assigned to share some of that experience with the RHW Team so that we will work together to chart our future direction.

Over the last year, our team adapted to many challenges. Among them, the coronavirus (COVID-19) pandemic and variant developments tested our resiliency and adaptable tools and procedures to continue operating in the office/labs, while both traveling and under telework environments. We have made great progress with how we we employ tools to advance our mission. In addition, the 2020 division re-organization/merger required adjustments to common processes and procedures. Moreover, with the Higher Headquarters (HHQ) strategic alignment and guidance, our team made sure to correct and fine-tune themselves where needed. We maneuvered accordingly while supporting our mission.

Our team contributed on many larger, multi-directorate/discipline, and transformational developments for the Warfighter. In addition, we met multiple technical and scientific milestones, adjusted to support the Air and Space Forces as one lab, published work, and participated in conferences. We also supported several communities of interest and exercises, demonstrated and transitioned products, strengthened our team, and cultivated our national and international relations with our partners in government, academia, and industry. Impressive year–great job!

As we welcome another year, we focus on our mission of "Enabling a more informed, agile, and lethal force by delivering revolutionary capabilities and enhancing preparation and mission execution across the full range of military operations." Our work centers heavily on enhancing the Warfighter's cognition in support of military operations: from the Warfighter's mission training, planning, execution, assessments, to supporting

system interfaces, sensing, analytics, and teaming (including with machines). The involvement in this field is wide, from knowledge acquisition to awareness and understanding. While we look to our contribution to the future forces, we continue to expand our connections to the Joint All Domain Ops and autonomy technologies.

Throughout the rest of this publication, you will find out we use two main core technical competencies (CTCs) for this field (Training and Adaptive Warfighter Interfaces) to provide scientific vision, guidance, and direction for basic and applied research. Our Training CTC focuses on enhancing our ability to acquire higher-quality and -quantities of knowledge. The Adaptive Warfighter Interfaces CTC explores the interface technologies, analytical decision aiding tools, and situationally-adaptive augmentation methods to integrate the warfighter and autonomous/intelligent technologies into maximally-collaborative warfighting teams.

We also employ three Product Lines (PL)—Readiness, Analytic Tools, and Airman Machine Integration—to engage with customers and serve as enterprise leaders and strategic managers of products (from definition through product transition). Our PLs link our CTCs efforts to Advance Technology Development of tools and/or products. The Readiness PL supports force readiness (training, proficiency and currency) for current and future operations. The Analytic Tools PL streamlines workflow and enables automation for data correlation opportunities while enhancing decision making. Lastly, the Airman Machine Integration PL aims for situationally-adaptive and -scalable interfaces and decision aiding tools for faster and more accurate situational awareness, decision making, and maximized collaborations and teaming (including distributed and with machines).

Our team is strong, talented, and ready for the challenges. I'm looking forward to the collaborations, great discoveries, advancements, and developments yet to come. Ready, Ready... Fight's-ON! ☆

Col Alfredo Rivera, Acting Division Chief, 711 HPW/RHW

> **Our work centers heavily on enhancing the Warfighter's cognition in support of military operations: from the Warfighter's mission training, planning, execution, assessment, to supporting system interfaces, sensing, analytics, and teaming..."**
>
> **— Col Alfredo Rivera**
> Acting Division Chief, 711 HPW/RHW

Photo by Master Sgt. Eric Harris

Dr. Winston "Wink" Bennett
Readiness Product Line Lead, 711 HPW/RHW

READINESS

A hearty welcome to the 2021 I/ITSEC Edition of *Fight's ON!* and the Readiness Product Line for this year; I'm thrilled to be back in person with you all! There is plenty to talk about this year and I am excited to share our research and field work as the Readiness Product Line Lead for our Division. The merger of two Divisions we wrote about last year continues to open new opportunities for collaboration and support.

There are plenty of positives that occurred this year to share with you all. First, I am elated to share that one of the concepts we discussed last year as a new cross-Wing and -MAJCOM effort, called Just In Time MultiMission Airmen/Warfighters (JITMMA/W), has progressed, and this summer we received word that we would be funded as a new start product in Readiness—that's phenomenal news for us. It's now a corporate product for AFRL so let me describe a little more about the concept.

The JITMMA/W concept started in March of 2020 as a collaboration between Air Combat Command (ACC), AF Futures, and the Office of the Secretary of Defense (OSD) in direct response to human performance enhancement technology needs from ops associated with Agile Combat Employment (ACE) and mission performance at austere, deployed locations. The effort includes integration and field demonstrations of the enabling technologies that underpin just-in-time training and task performance support. The end-state will enable Airmen and Guardians to work across mission tasks through training, performance support and resilience monitoring, and intervention to seamlessly adapt to changing operational needs and provide a broader range of mission support in austere, deployed environments. We're planning iterative demonstrations that start in our virtual and constructive model-based systems within engineering-based environments, with Subject Matter Experts (SMEs) and engineers evaluating the alternative capabilities and their prospective mission impacts first. We'll then take the best of this digital engineering-based approach to actual technology integration and evaluation for a series of live field tests.

This year, we have made stronger connections to the field events planned for ACE and developed some notably strong collaborations and connections to the Technical Training Transformation (T3) in our Air Education and Training Command (AETC). We have also started reaching out to our other Services for collaboration on the concept from their needs and perspectives. We are planning on an initial integrated demo of some of the key-technology capabilities in the spring of 2022. At the same time, we are reaching out to industry and academia with a series of targeted small business innovation research (SBIR) and technology transfer and research (STTR) opportunities to further innovation in both the concept and capabilities—very exciting stuff!

This past year also saw us complete the stand up of our Content Data Standards Big Data/Data Lake effort. Last December, we hosted a major AFWERX challenge virtual workshop. That workshop, the expertise, and recommendations we were able to draw on from industry, academia, and other government agencies was foundational to our stand up. We're leveraging innovations coming from the challenge workshop to standardize how we capture, store, retrieve, and package our research data for a variety of applications, such

as machine learning, mathematical modeling, and longitudinal trend analyses of training and readiness effects. A major push for the Data Lake is related to codifying the past 20 years of training data that we continue to harvest and the migration of some tools into ops to do the same on a routine basis. In fact, we are integrating several of our tools into operational mission training center software integration labs today as a means of de-risking those technologies supporting the data capture and feedback in the operational mission training centers and ranges in the next year or two. Here too, we have developed some innovation opportunities in collaboration with AFWERX for SBIR and STTR opportunities in big data and analytics associated with large-scale data repositories.

Recently, we've finished our first full year of a new multiyear, multidisciplinary effort to conduct research, development, and integration of science and technologies supporting Headquarters Air Force and Air Combat Command. The work is centered on several areas of need from the ops community and provides research and engineering support for those areas. Examples include creating rule sets for coalition, distributed mission operations (DMO) interoperability for tactical training, and validating specifications for performance data and encryption of the data from live and simulator sources. It also supports 4th- and 5th-gen integrated training, using the testbeds we have in the Division; subsequently, it also integrates performance measurement and feedback technologies identified to support the CAF Future Training Concept into selected mission training center software integration labs. There is also support here for creating and transitioning live, virtual, and constructive technologies and tools into ops training and rehearsal.

We're continuing the two efforts that were driven by General Mike Holmes during his tenure as Commander of ACC (COMACC). The first related to this effort is developing a web-based repository that will allow USAF training content developers, instructional designers, and instructors to share and reuse critical AR/VR-related technologies (e.g., reusable models, source code, executable code, APIs, SDKs, performance measures and dashboards, etc.) with one another. We've continued to gather data virtually and are now heading to the field to meet with innovators and developers directly to capture more specific details and applications of the technologies in research and ops. This will provide deeper details for tracking, integration, and fielding of AR/VR-based solutions more quickly and cost-effectively by not having to "reinvent the wheel" every time a new application is developed or someone looks for a policy precedence for software and hardware accreditation, or certifications and related lessons learned (and other data artifacts).

The progress that continues on the second of these is a collaboration with researchers at the Institute for Defense

Drives "innovation at the speed of Ops"

Advances the state-of-the art in learning, performance and modeling

Develops solutions and capabilities—informing Air Force vision and investment

Aligns with academia, acquisition, and industry to transition effective and efficient methods and technologies

Contributes to the operational readiness of our Airmen and Guardians

Analyses. This effort has developed an "effectiveness framework for AR/VR training" based on science and empirical findings. It is also documenting current research and applications of the technologies in military education and training. This past year, we continued qualitative and quantitative reviews of the published literature to make recommendations for research and practice.

Finally, our academic, industry, and international partners continue to grow their interest and involvement in collaborations with us. We have several new agreements, partnerships, and contract activities underway. Our Division is growing its already-strong support of our international standards development and involvement in multinational research and technical groups.

I invite you to contact me or any of our team members to explore opportunities for collaboration and partnership. ☆

Dr. Winston "Wink" Bennett, Readiness Product Line Lead, 711 HPW/RHW

Dr. Deirdre Mahle
Airman-Machine Integration Lead,
711 HPW/RHW

> **"The ingenuity, creativity, and intuitiveness of the human brain, coupled with the speed and capacity of AI/ML systems, generating accelerated and resilient decision-making in all domains is truly what sets us apart from the adversary."**
>
> **— Dr. Deirdre Mahle**
> Airman-Machine Integration Lead,
> 711 HPW/RHW

AIRMAN-MACHINE INTEGRATION (AMI)

Wow, I can't believe it has been a year since I joined the Warfighter Interactions and Readiness Division as the Airman-Machine Integration Product Line Lead (AMI PLL). It's been an exciting and challenging year to say the least. Despite the ongoing pandemic, we have still managed to deep dive into the different technologies and efforts used to create and enhance human-machine teams– and I have to say it has been mind-blowing! The AMI PL is delivering important products, such as a library of user interfaces, including the Pilot Vehicle Interface controls tailored for swarm Small Unmanned Aircraft Systems (SUAS) being transitioned to the Skyborg Vanguard. A distributed play-calling and dynamic course-of-action, real-time analytic toolkit is currently integrating with the Medusa multi-domain C2 system for base defense, allowing for control of multiple unmanned systems in multiple domains. The ingenuity, creativity, and intuitiveness of the human brain, coupled with the speed and capacity of AI/ML systems, generating accelerated and resilient decision-making in all domains is truly what sets us apart from the adversary.

In the past year, as our researchers have developed robust operator interfaces and advanced, situationally-adaptive interface technologies and decision aiding tools, our focus has shifted to supporting the Advanced Battle Management System (ABMS) as an enabler for Joint All-Domain Command Control (JADC2). Distributed teams are not a challenge to overcome but an advantage to be exploited. Distributed also means agile, but only if we are seamlessly synergizing human cognitive capability with intelligent machine teaming, thereby maximizing the performance of both. Our research teams are also connecting across the AFRL on several key topic areas for transformational, game-changing technologies presented at the 2nd summit of technical, operational, acquisition, and planning communities (WARTECH 2.0), as well as Transformational Capabilities Office (TCO)-supported explore campaigns for future force capability priorities. Mayhem, the multi-mission, hypersonic strike weapon, is a high-priority topic area that will rely on state-of-the-art Mission Planning and Debrief, augmented with AI in order to deliver the unique characteristics of a hypersonic asset. High Speed Vertical Take Off and Landing (HSVTOL) is a key capability for runway independence and will close gaps for personnel recovery, logistics under attack, and medevac missions. Our team is providing roadmap inputs for Human System Integration, Human Machine Integration, and Human Factors for crew cabin design requirements, as well as protections for a recovered, isolated personnel. 2022 looks to be an even more exciting year, and I can't wait to see what else is in store for AMI! ☆

Dr. Deirdre Mahle, Airman-Machine Integration Lead, 711 HPW/RHW

ANALYTIC TOOLS

The contention that "Information Dominance" or "Decision Superiority" is the key to success in conducting JADC2 in a GPC environment may hold some truth, but I assert that the advantage goes to those with "the will to act." As an analog, it could be argued that most people have the information and knowledge to improve their health via proper nutrition, physical activity, recovery, and relaxation. However, data suggests that the overwhelming majority of the population does not implement this knowledge, despite knowing the consequences. Data, information, knowledge, and understanding alone is insufficient to affect change.

How does the RHW and the Analytic Tools (AT) Product Line (PL) close the gap between information and action? Our paradigm is inherently human-centric, which means our ideal research/engineering approach is to place human cognition in the center and build the technology around it. In this vein, the mission of the AT PL is to:

1. Develop tools that enable humans to effectively process information.
2. Integrate those tools into processes that transform information into action.

In FY22, the AT PL will re-purpose a combat-proven simulation environment that's used to train Special Operations Forces into a Live-Virtual-Constructive (LVC) Laboratory that is extensible to conventional-led JADO/JADC2 operations in a GPC environment. This LVC Lab is the 21st Century's "digital engineering" equivalent of traditional mechanical and electrical R&D shops. Instead of measuring signals and bending metal, the purpose of the LVC Lab is to serve as:

1. An AFRL springboard for developing, testing, and transitioning cognitive-based analytic tools that are vetted by thousands of operational end-users.
2. A DoD "digital engineering" and research platform for designing and testing virtual solutions to complex problems, such as JADO/JADC2 and GPC.
3. A venue to collect "data as a product" that enables researchers across AFRL to understand and characterize human cognition and performance.

The LVC Lab is extensible to all platforms, payloads, and infrastructures in the space, air, ground, maritime, and cyber domains.

All interested parties are invited to contact me to discuss Analytic Tools, or other capabilities, whose goal is to transform information into action. ☆

Mr. Greg Beister, Analytic Tools Product Line Lead, 711 HPW/RHW

Mr. Greg Beister
Analytic Tools Lead, 711 HPW/RHW

> **The contention that 'Information Dominance' or 'Decision Superiority' is the key to success in conducting JADC2 in a GPC environment may hold some truth, but I assert that the advantage goes to those with 'the will to act.'"**
>
> **— Mr. Greg Beister**
> Analytic Tools Product Line Lead,
> 711 HPW/RHW

TRAINING

Dr. Glenn Gunzelmann
Training Core Technical Competency Lead,
711 HPW/RHW

"THE WAYS IN WHICH WE TEST, EVALUATE, AND TRAIN WITH [EMERGING TECHNOLOGIES] DO NOT MEET CURRENT OR FUTURE DEMANDS."

— Gen. Charles Brown
Air Force Chief of Staff

MULTI-CAPABLE READINESS:

Exquisite Expertise Meets Agility for Multi-Mission Operations

Welcome once again to *Fight's ON!* What a year it has been. I started my introduction last year and in 2019 with Heraclitus – "The only constant is change." This year, I am reminded of another quote – "The more things change, the more they stay the same" (Jean-Baptiste Alphonse Karr, 1849). Change continues to roll upon us at a breathtaking pace. However, through that change is remarkable stability in the message regarding the future forces and critical role of training for ensuring mission effectiveness. In future operations, the readiness and effectiveness of our people will be the differentiator, and we must ensure that our training strategies, technologies, and practices keep pace with the rapidly-changing operational environments and the world that we are preparing our Airmen and Guardians for.

The quote from the Air Force Chief of Staff to the left reinforces the need for transformation in training for the United States Air and Space Forces to meet that need. We must enable tomorrow's Airmen and Guardians by matching our readiness ecosystem to the requirements and opportunities created by new technologies combined with an increasingly complex and dynamic operational environment–this is no small task. The Training Core Technical Competency (CTC) within the Air Force Research Laboratory (AFRL) leads the Air and Space Forces into that future by creating, demonstrating, and delivering revolutionary training and readiness capabilities at the speed of operations. We must ensure that our Airmen and Guardians have the exquisite expertise to outperform our adversaries in every domain of conflict and the agility and adaptability to support multi-mission operations whenever, wherever, and whatever the need(s).

THE TRAINING CTC CONTINUES TO PURSUE THIS MISSION THROUGH THE PURSUIT OF THREE INTERRELATED GOALS:

Personalized, Proficiency-Based Readiness

The core of the revolution of training and readiness we pursue is the shift from a calendar-based, industrial age model to one that focuses on individuals and mission effectiveness. We must advance our ability to measure and assess performance and use that information to make quantitative decisions about every individual's capacity to effectively execute missions in complex, uncertain operational environments. This capability must scale to support the readiness of every Airman and Guardian.

Global, Persistent, Live-Virtual-Constructive (LVC) Training

Providing the training necessary to ensure personalized readiness requires a readiness ecosystem that can deliver training and instruction whenever and wherever it is needed. Live-Virtual-Constructive training has been a core concept in the arsenal for maintaining readiness for both the Air and Space Forces for many years. However, there are still critical gaps in the technology for delivering tailored training experiences that seamlessly integrate live, virtual, and constructive entities to provide the kind of full-spectrum training required to ensure mission effectiveness.

Team Training for Humans and Machines

The deeply-interconnected nature of modern-military operations means that we must go beyond individual proficiency by measuring and assessing the capacity of teams to work in an integrated and seamless manner to achieve mission objectives. Our training ecosystem must also take into account rapid advances in the applications of artificial intelligence to create autonomous systems, leading us toward machine systems that will increasingly influence team dynamics and performance. Further, training environments must provide opportunities for humans and machines to practice and learn together – mutually adapting to individual idiosyncrasies, environments, and operational requirements to optimize performance and mission effectiveness.

These are audacious goals that require advances in training technology and changes in training practice, along with a new culture of readiness that emphasizes quantitative assessments, predictive validity, and agility to meet uncertain operational demands. We are pursuing research efforts on all of these fronts, working with universities, industry and international partners, other government agencies, and operational stakeholders to leverage the state-of-the-art in science and technology while remaining attentive to operational requirements and constraints.

Under new leadership at the Core Research Area (CRA) level, we have refined our research strategy, streamlined our research efforts, and targeted our investments to establish critical capabilities to enable transitions through the organization's product lines, engagements with program offices at the Air Force Life Cycle Management Center and elsewhere, and through direct engagements with operational customers.

Take some time to read about our diverse, innovative, and impactful research efforts in this issue of *Fight's ON!* Even in the face of persistent, dramatic change, AFRL's Training CTC continues to deliver on its vision to be the Department of Defense's premier science and technology organization for training and readiness. ☆

Dr. Glenn Gunzelmann, Training Core Technical Competency Lead, 711 HPW/RHW

COGNITIVE MODELS (CM)

Dr. Megan Morris
Cognitive Models Core Research Area
Lead, 711 HPW/RHWM

TWO PROJECT AREAS

- **Teachable Models for Training**
- **Multiscale Models for Cognitive Performance**

The Cognitive Models Core Research Area (CRA) advances the state-of-the-art in the computational and cognitive sciences to develop technologies for on-demand training support and continuous readiness monitoring. Specifically, we are conducting leading-edge foundational and applied research on human cognitive capacities, developing adaptive and individualized modeling capabilities, and conducting novel integrations with other technological approaches to develop synthetic agents and effective cognitive performance and prediction systems.

Two technical challenge areas (lines of effort (LOE)) drive our CRA work. The first challenge area is Teachable Models for Training (TM4T), which focuses on how we can improve mechanisms for learning and knowledge representation in software models of human cognition to increase factors such as affordability, adaptability, generalization, and personalization of these models. To meet this challenge, we are developing models that can detect, identify, and resolve knowledge gaps within the model, adapt to dynamic and personalized communication, and rapidly track and train across various skills for tailored training agent needs. These capacities enable rapid generation of high-cognitive-fidelity models to support personalized training as synthetic teammates, adversaries, and instructors.

The second challenge area is Multiscale Models for Cognitive Performance (MSM), which focuses on how we can integrate models of different levels of analysis–physiology, neuroscience, cognition, and behavior–to provide converging evidence about current and future mission effectiveness based on various stressors. To meet this challenge, we are developing novel techniques that focus on integrating these levels in an individualized manner. The goal is to provide real-time, behavior-specific assessments and predictions of readiness based on stressors such as fatigue, sustained attention, workload, pharmaceuticals, toxins, and environmental factors. The following articles highlight research in these areas. ☆

Dr. Megan Morris, Cognitive Models CRA, 711 HPW/RHWM

TEACHABLE MODELS FOR TRAINING

The development of intelligent systems capable of operating as team members, adversaries, or tutors within team training scenarios is costly and time consuming. A consistent issue is the requirement for developers to analyze, organize, and integrate task knowledge within a modeling-and-simulation framework (a.k.a. the knowledge engineering bottleneck). To widen the knowledge engineering bottleneck for greater efficiency in intelligent agent development, we are leveraging our understanding of human cognition formalized in executable models that can learn from instruction. Three related tasks are focused on making progress to achieve the goal of teachable models.

TEACHABLE MODELS FOR TRAINING TASKS

Knowledge Gap Detection, Identification & Resolution

Adaptive Linguistics

Rapid Multi-Skill Acquisition in Models

Each task is covered in greater detail in the following pages. ☆

Dr. Christopher Myers, Senior Cognitive Scientist, 711 HPW/RHWM

TEACHABLE MODELS FOR TRAINING OVERVIEW

The nature of the research and development within Teachable Models for Training requires a multidisciplinary approach. Across the three tasks, contributions are provided from computer science, machine learning, cognitive science, computer science, and modeling and simulation.

Graphic by 711 HPW/RHW

Graphics by 711 HPW/RHW

KNOWLEDGE GAP DETECTION, IDENTIFICATION AND RESOLUTION

As the ability for machines to learn from instruction and experience increases, their ability to identify and resolve gaps in their newly acquired knowledge will become increasingly important. The ability to detect, identify, and resolve gaps in knowledge required to successfully execute newly instructed missions will help to reduce autonomy failures and mistakes. Both the research and development(s) adopt the approach that it is preferable to identify inconsistencies and absences within a knowledge base early rather than to assume that a system's knowledge is correct and complete. To this end, the ability to detect knowledge gaps is critical. Informally, knowledge gaps arise when the information or capabilities available to an agent or autonomous system are inconsistent or missing. This leads to inefficiency in performing a task or failure to perform the task at all. Knowledge gap detection, then, is the ability to recognize when there are inconsistencies in knowledge that can cause a knowledge gap to arise.

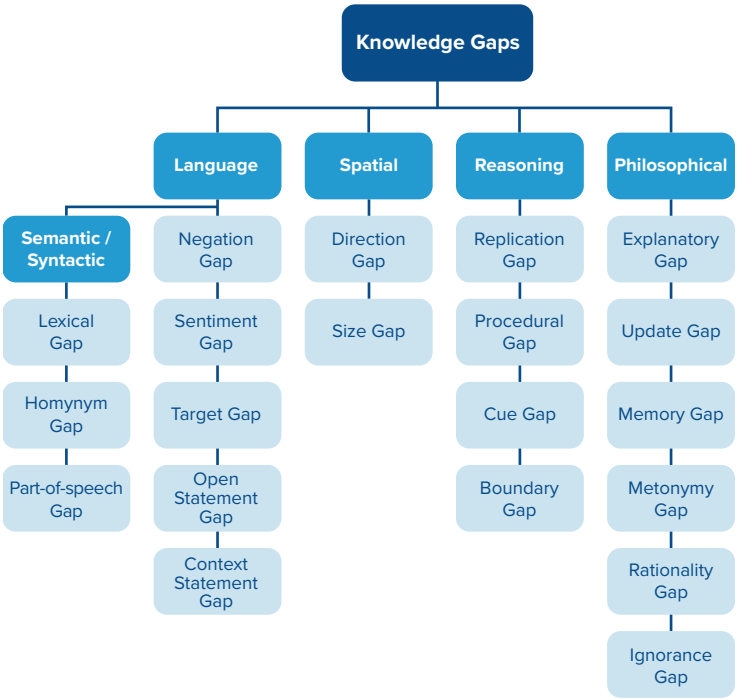
Knowledge gap identification consists of the ability to classify the gap as belonging to a certain category of knowledge gaps. Once a gap is identified, it becomes possible to resolve it through knowledge gap resolution, either automated or with a human trainer's assistance.

The team is investigating the ability to perform knowledge gap detection and resolution within two prominent approaches to agent development:

- Symbolic/hybrid computational cognitive architectures
- Artificial neural networks

Further, two domains are being leveraged for knowledge gap detection and resolution:

- Instruction learning
- Visual question answering



Through an Educational Partnership Agreement with the Ohio State University Department of Computer Science & Engineering, approaches to leveraging visual question-answering tasks as a means to test, evaluate, verify, and validate developed processes and mechanisms associated with knowledge gap detection and resolution (Bajaj, et al, 2020; 2021) have been investigated. ☆

Dr. Christopher Myers, Senior Cognitive Scientist, 711 HPW/RHWM

Dr. Christopher Stevens, Research Psychologist, 711 HPW/RHWM

Mr. Michael Dougherty, Computer Scientist, 711 HPW/RHWM

Ms. Mary Freiman, Contractor, Aptima

Dr. Brent Fegley, Contractor, Aptima

Dr. Srini Parthasarathy, Professor in Computer Science, The Ohio State University

ADAPTIVE LINGUISTICS

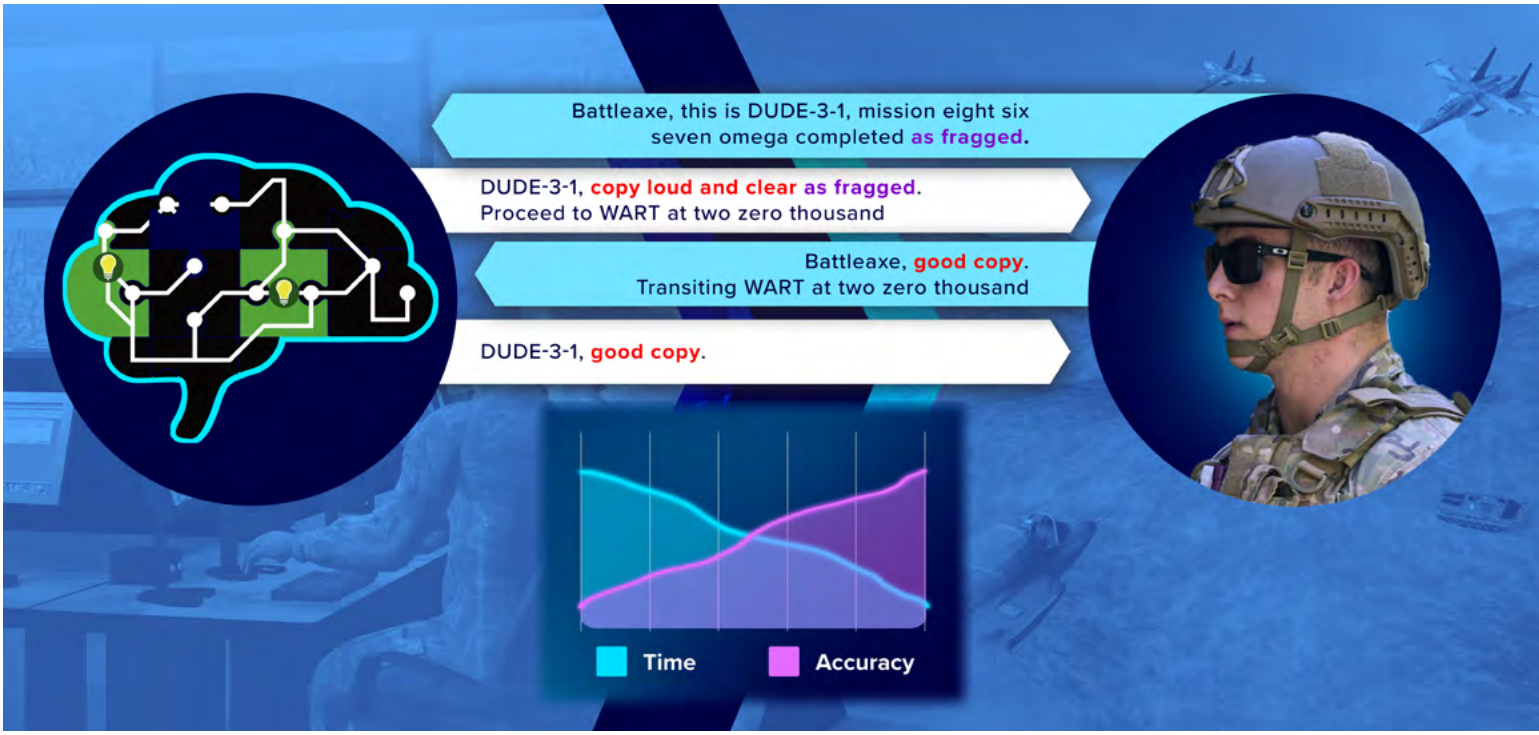
For decades, research on human-computer interaction was focused on identifying barriers and paths to optimal human-machine performance. With the increasing capabilities within intelligent machines, human-machine teaming has emerged as a critical area for advancing Department of Defense (DoD) goals. The ability to seamlessly team with intelligent machines crucially requires advancements in machines' capabilities to adapt to changes in communications. The Adaptive Linguistics task focus on how to expand the language capabilities of current cognitive models and synthetic teammates to include the ability of incorporating novel linguistic items or novel uses of known linguistic items.

Further, the research efforts are addressing how to enable machines to tailor language to particular task contexts and teammates. This task includes novel empirical investigations with all-human teams to achieve foundational knowledge in how humans adapt language usage to each other in order to accomplish a particular task; additionally, it leverages high-fidelity, mission-relevant testbeds (e.g., Mission Planning and Debrief) as a way to develop and test cognitive models of adaptive, task-based language use. Expected results include identifying the best candidate linguistic structures for cognitive models and synthetic teammates to adapt based on which structure(s) humans adapt and contribute to improved team performance. Impacts for the Warfighter include more naturalistic interactions with synthetic teammates and intelligent tutors, less cognitive load from needing to learn and apply restrictive communications with synthetic teammates, and improved coordination between human teammates and synthetic systems.

This past year, we determined that developing cognitive models and synthetic teammates with adaptive language capabilities required expanding current language representations and processes along with innovating novel approaches to represent linguistic content and structure. To overcome this gap, we submitted and were successfully awarded a $70K seedling from the 711 HPW Chief Scientist's Office to develop such processes to be tested and validated in a cognitive model paired with human teammates in cooperative-training settings. The team is currently pursuing several implementation options for the cognitive model while we await human-human team data from a cooperative card game task—which has been delayed due to COVID-19. ☆

Dr. Sarah Bibyk, Research Psychologist, 711 HPW/RHWM

Graphic by 711 HPW/RHW



RAPID MULTI-SKILL ACQUISITION IN MODELS

Imagine the following scenario: A training exercise will take place and your unit is able and ready to execute the mission. Unfortunately, most of your teammates are in quarantine due to COVID-19 and can't participate. As a result, intelligent agents will serve as surrogate humans, filling in for your absent teammates and playing their roles in non-perfect, human-like ways. These agents will be tailored to possess varying levels of performance proficiency across key mission skills that are aligned with the individual differences inherent to humans. This type of proxy human realism is required to optimize the learning outcomes for human trainees under plausible, and namely, less than perfect circumstances. To achieve this feat, agents must possess the ability to track performance within and across targeted skills and be rapidly tailored to achieve desired proficiency levels.

RHW researchers are developing a new line of research to understand and evaluate potential tradeoffs and cascades of implications involved in multi-skill acquisition, ultimately identifying and prescribing agent training regimens across a larger set of desired skills/competencies. This work leverages a seedling award from the 711 HPW Chief Scientist's Office, demonstrating the integrated-value of cognitive models and machine learning algorithms enmeshed in an ensemble approach. Through the use of ensemble models, both enhanced predictive validity and minimized data requirements have been revealed. This new research will bring to bear the ensemble modeling formalism to:

1. Rapidly train targeted skills to desired levels of model proficiency.
2. Track learning and forgetting across multiple skills in cognitively-plausible ways.
3. Assess the impact of one skill on another considering additional learning gains or decay of other skills.

This work will afford the ability for skills to be rapidly trained for synthetic teammate use; allow for the evaluation of the training impacts to be quantified, as they relate to other required skills; and serve the development of intelligent tutoring systems aimed at optimizing the interleaving of skills for enhanced acquisition and sustainment across the full set of desired skills of interest. ☆

Dr. Tiffany Jastrzembski, Senior Cognitive Scientist, 711HPW/RHWS
Ms. Olivia Leung, Research Computer Scientist, 711HPW/RHWM

Graphic by 711 HPW/RHW

RESEARCH HIGHLIGHT

COGNITIVE FIDELITY IMPROVEMENTS TO AGENTS IN ADVANCED FRAMEWORK FOR SIMULATION, INTEGRATION, & MODELING

A collaboration between the Airman Systems (RH) and Aerospace Systems (RQ) directorates within AFRL has been researching and developing improvements to the cognitive fidelity of agents within the Advanced Framework for Simulation, Integration, and Modeling (AFSIM). Currently, ASFIM agents are unconstrained in their information processing capabilities– they are capable of using all information instantly to determine which action to take from all possible actions. AFSIM users and developers alike have identified the superhuman capabilities of AFSIM agent information processing as a detriment to the accuracy of simulation outcomes as the agents do not accurately reflect human cognitive capacities.

The RH-RQ team has focused research and development on limiting agent capabilities within the Observe and Orient components of the Observe-Orient-Decide-Act (OODA) loop; specifically, limiting the ability for agents to rapidly and completely process information for all tracks available to any particular agent with capabilities observed in humans. Further, the team is working toward the integration of directed-energy effects on cognitive performance, specifically the effects of laser dazzle and flash blindness on perceptual encoding abilities while piloting aircraft. The research is being conducted in a High-Value Aerial Asset scenario within the AFSIM architecture by using varying levels of complexity to evaluate when and how agents' increased cognitive fidelity affect the simulation outcomes. ☆

Dr. Christopher Myers, Senior Cognitive Scientist, 711 HPW/RHWM
Dr. Christopher Stevens, Research Psychologist, 711 HPW/RHWM
Dr. Kristen Liggett, Principal Human Factors Engineer, 711 HPW/RHWCI
Dr. Aaron Hoffman, Physicist, 711 HPW/RHD
Dr. Nick Hanlon, Aerospace Engineer, 711 HPW/RQS

Although the effects of laser dazzle are non-lethal, it has the ability to temporarily disable its targets (in this case, pilots) with flash blindness.

Further, the team is working toward the integration of directed-energy effects on cognitive performance, specifically, the effects of laser dazzle and flash blindness on perceptual encoding abilities while piloting aircraft."

— Dr. Christopher Myers,
Senior Cognitive Scientist,
711 HPW/RHWM

Photo by U.S. Airman Shawna Keyes

MULTISCALE MODELS FOR COGNITIVE PERFORMANCE

Warfighters face a variety of operational and environmental stressors as they perform their duties. The goal of the Multiscale Models team is to combine models at different analysis levels (physiological, cognitive, and behavioral) to predict the effects of stressors on Warfighter readiness and cognitive performance. Specifically, we integrate high-fidelity models that represent the physical and biological effects of stressors, such as fatigue, caffeine, toluene, hypoxia, and workload, with cognitive models that represent moment-to-moment cognitive activity. The physiological models modulate the mechanisms in the cognitive models, producing a performance prediction. Work within this line of effort is divided into three tasks, each focused on a different type of cognitive stressor:

1. Fatigue and Sustained Attention Performance Impacts
2. Individualized Cognitive Load Profiling
3. Physiocognitive Modeling of Environmental Stressors

These integrated capabilities will improve risk assessment and mitigation with respect to Warfighter cognitive performance. ☆

Dr. Christopher Stevens, Research Psychologist, 711 HPW/RHWM

PHARMACEUTICALS

HYPOXIA

TOXINS

WORKLOAD

FATIGUE

Graphic by Mr. Will Graver

FATIGUE AND SUSTAINED ATTENTION PERFORMANCE IMPACTS

The Fatigue and Sustained Attention Performance Impacts task focuses on advancing modeling of human fatigue and sustained attention in operational settings to increase the efficacy and efficiency of fatigue risk management (FRM) tools in the Department of the Air Force. Specifically, we focus on developing models that produce individualized, real-time, behavior-specific estimates of cognitive performance to inform mission readiness.

It has been a busy year, with lots of progress on multiple fronts. For example, our team had great success earlier this year providing data analysis and interpretation support for a study led by the 341st Missile Wing at Malmstrom Air Force Base (AFB) that explored alternative schedules for manning Missile Alert Facilities. Based on the results, the 20th Air Force has adopted the schedule permanently, marking the first such change in this domain. It is estimated this will improve medical readiness by 74.6%, increase morale for the missileer career field (13N), and generate fuel savings of $75,000/year.

A report on our study with C-17 aircrew at Travis AFB is available at the Defense Technical Information Center. The study examined self-reported and objective fatigue in aircrew and compared prescriptive sleep schedules from a FRM tool to actual aircrew sleep schedules during missions. Results suggested the aircrew are concerned about fatigue in the community and that the tool underestimates fatigue during missions, leading to increased safety risk.

Additionally, we are currently wrapping up data analysis for a study with Airborne Warning and Control System crews. The study aims to identify current fatigue risks and challenges in the community and examine possible alternative sleep schedules to reduce fatigue.

Futher, our team continues to enhance our mobile fatigue application with new capabilities. Currently, we're integrating sleep estimates from physiological devices, such as the Oura Ring and Garmin fitness watches, to produce fatigue assessments and predictions. We recently presented work at the International Conference on Cognitive Modelling on an effort to individualize the embedded biomathematical fatigue model in the application with real-time performance data from the Psychomotor Vigilance Test. The mobile fatigue application work supports a broader technology demonstration project within the 711th Human Performance Wing (711 HPW) called Fatigue Optimized Cognition Under Stress (FOCUS).

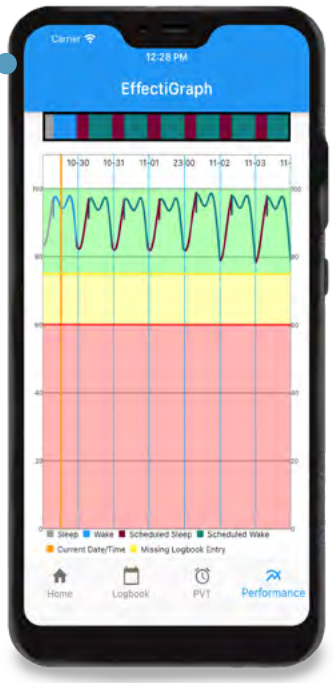
As well, we have teamed up with researchers from Johns Hopkins University (JHU) to support integrating fatigue modeling into a wargaming logistics simulation, known as iSWAT, to highlight human performance impacts on personnel availability and effectiveness. Our team is developing an application that provides fatigue estimates for aircrew and maintenance personnel given simulated scheduling from the simulation software. The JHU team is focusing on alternative fatigue modeling implementations and incorporating effects of caffeine and sleep aids on fatigue assessments and predictions. Once again, this effort is part of a broader 711 HPW project–called Human Digital Twin–which is exploring capabilities for representing human biological, physical, and cognitive characteristics at a level of fidelity that provides useful constraints for system design, readiness and performance monitoring, and planning.

Finally, we will conduct a study with an airlift wing at Yokota Air Base to develop an approach for including fatigue estimates from a smart wearable strap into the Wing's operational risk management worksheet in order to increase aircrew performance and mission safety. We also plan to conduct a study with fighter pilot samples from several bases where we will examine fighter pilot fatigue in non-combat operations. We will collect physiological data from personal fitness devices to derive objective fatigue values during operations and integrate this data into our mobile fatigue application to provide real-time, individualized, objective situational awareness of fatigue for the fighter pilot community.

Additionally, we plan to conduct a collaborative project with both the United States Naval Surface Force at Naval Amphibious Base Coronado and the Naval Medical Research Unit Dayton to develop a cognitive modeling capability aimed at predicting individualized performance degradations in operations due to fatigue and workload across different military domains. Specifically, we will conduct a series of simulator studies with C-17 mobility aircrew and Naval bridge crew and use that data to develop the modeling capability. ☆

Dr. Megan Morris, Research Psychologist, 711 HPW/RHWM

To the right is a screengrab of the Android/iOS mobile fatigue application prototype that will be capable of taking in user inputs and habits, exploiting these modeling advancements, and producing individualized fatigue assessments and predictions for the specific Airman.



Graphic by Ms. Shania Horner

INDIVIDUALIZED COGNITIVE LOAD PROFILING

Warfighters face demanding, complex, and ever-changing operational tasks. In the Individualized Cognitive Load Profiling task, we combine state-of-the art machine learning techniques with cognitive models to use converging evidence from physiological, behavioral, and subjective measures of cognitive workload to classify a Warfighter's workload level and predict their performance. We use information coming from multimodal physiological channels, such as heart rate and electroencephalography (EEG), to classify at a high-level whether an individual is experiencing high workload. To make this classification, we leverage artificial neural networks (ANNs) and Hidden Semi-Markov models. We then examine the activity in cognitive models of a task to determine which cognitive capacities (e.g. memory and visual perception) are likely to be overburdened.

Currently, we are creating models in four domains: Intelligence, Surveillance, and Reconnaissance (ISR; Fisher, Frame, & Stevens 2021), aerial refueling (Stevens, Fisher, & Morris, 2021), command and control (Stevens, Morris, Fisher, and Myers, 2019), and multitasking in aircraft-related tasks (Multi-Attribute Task Battery; Comstock & Arnegard, 1992). Previously, we demonstrated agreement between our cognitive models and estimates of global workload obtained from physiological metrics (Stevens, Morris, Fisher, Myers, 2019) and subjective reports (Stevens, et al., 2019; Stevens et al., 2021). In these tasks, we also plan to examine and predict the effects of individual differences, such as working memory capacity and personality traits on workload.

Moreover, we have plans in the near future to conduct an empirical investigation of the interactive effects of workload and fatigue (e.g., how does the level of cognitive effort that an individual exerts affect their rate of fatigue?) These models will provide important information about Warfighter tasking decisions, such as whether additional tasks can be added without sacrificing performance. Also, because they specify the individual cognitive capacities affected by workload, the models will provide information about when and how to mitigate high workload levels.

Our models can inform the technologies that mitigate risks associated with high workload, such as autonomous managers. Autonomous managers are software algorithms that assign tasks to members of a human-machine team to optimize performance based on each member's workload and expected performance (Frame et al., 2019). Through a series of simulations, we have shown that cognitive models combined with autonomous managers are a plausible method for improving workload and performance among a human-machine agent team in a laboratory ISR task (Fisher et al., 2021). We have plans to collect data to validate the results of these simulations through September-December 2021. ☆

Dr. Christopher Stevens, Research Psychologist, 711 HPW/RHWM

Dr. Megan Morris, Cognitive Models CRA, 711 HPW/RHWM

Dr. Leslie Blaha, Senior Research Psychologist and CMU Operating Location Lead, RHWL and Carnegie Mellon University

Ms. Olivia McCormick, Research Biomedical Engineer, 711 HPW/RHWM

Mr. Michael Dougherty, Computer Scientist, 711 HPW/RHWM

Mr. Justin Estepp, Research Biomedical Engineer, 711 HPW/RHBCN

Dr. Christopher Fisher, Cognitive Data Scientist, Cubic Global Defense

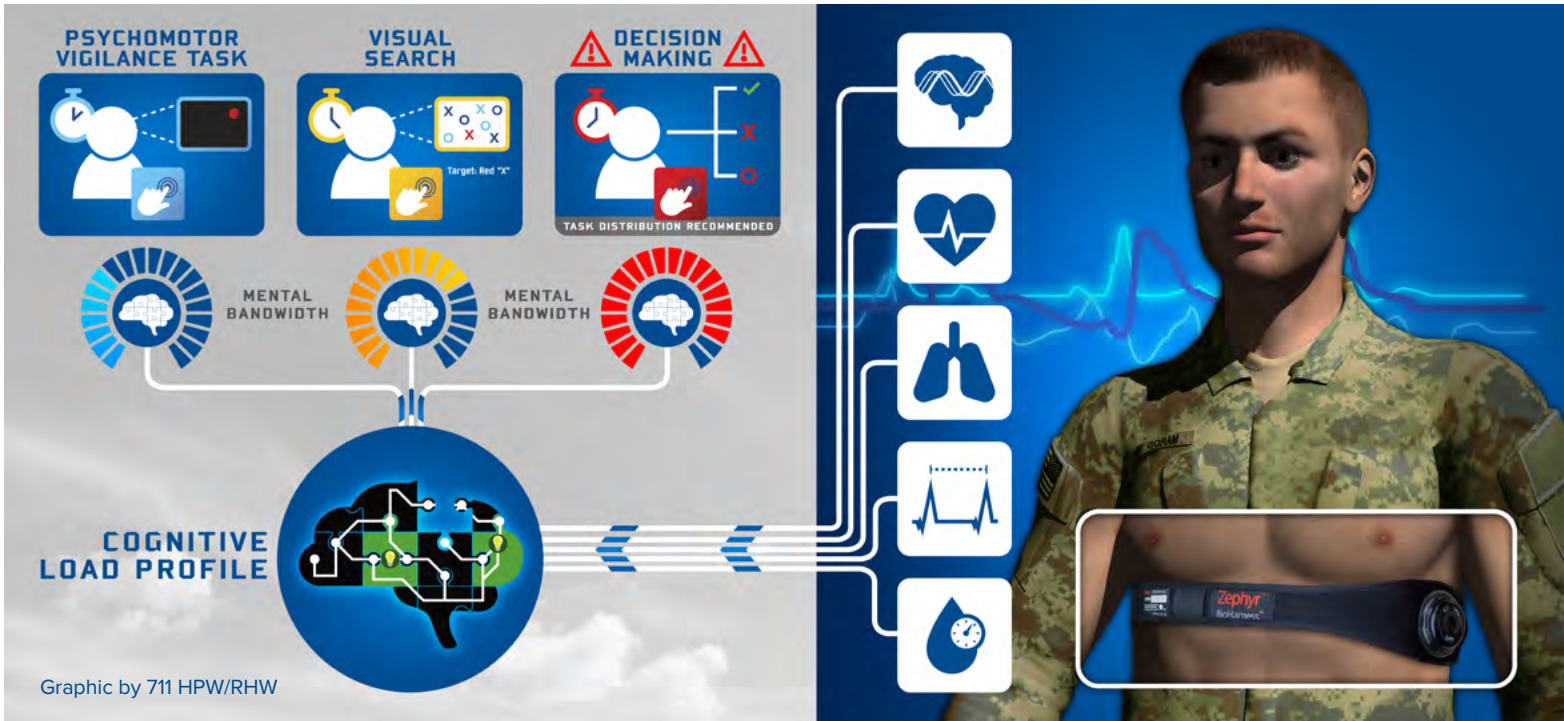
Dr. Garrett Swan, Cognitive Modeler, Cubic Global Defense

Dr. Mary Frame, Senior Research Psychologist, Parallax

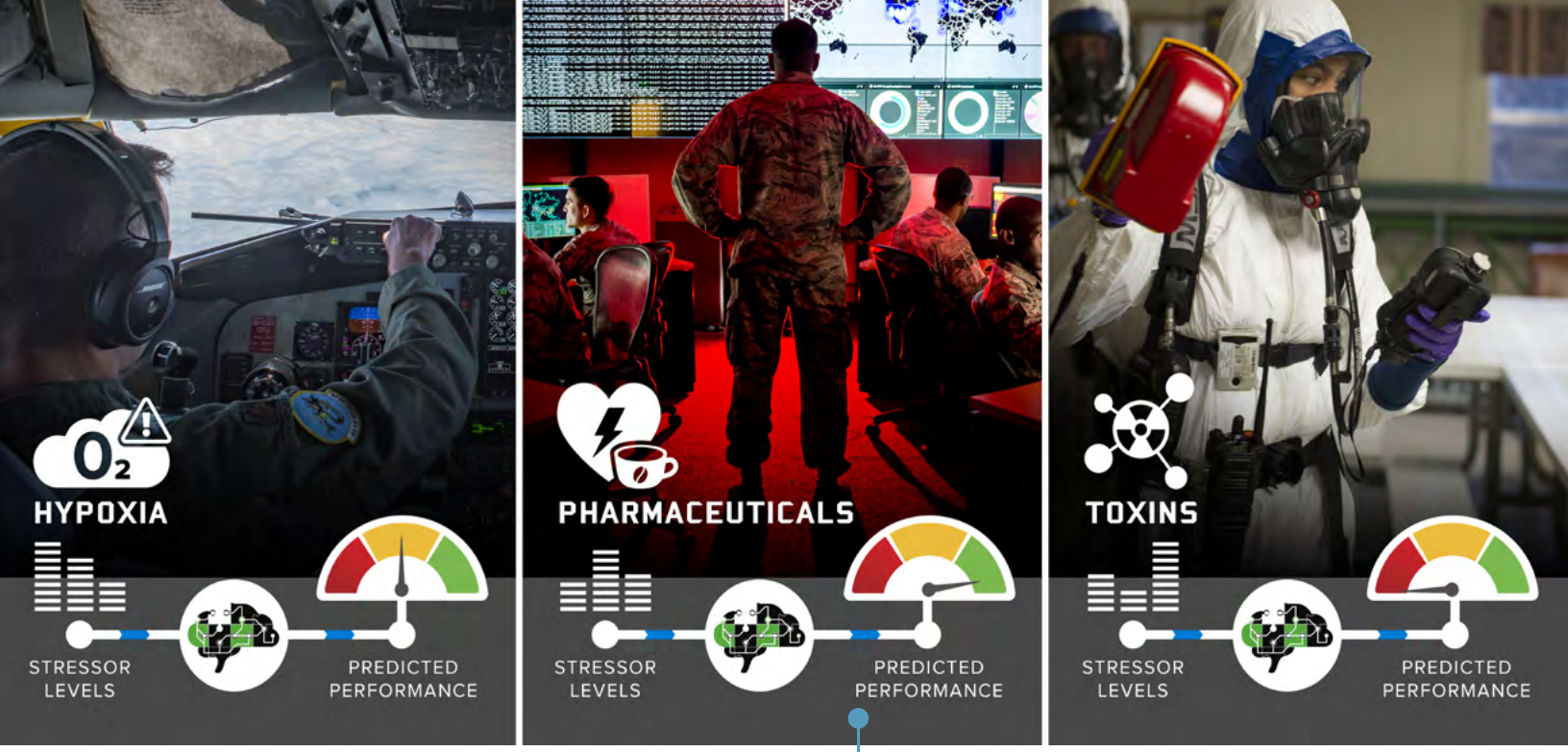
Dr. Molly Liechty, Research Scientist, Leidos

Ms. Samantha Klosterman, Biomedical Engineer, Ball Aerospace

Ms. Ashley Haubert, Associate Cognitive Research Coordinator, University of Dayton Research Institute



Graphic by 711 HPW/RHW



Graphic by 711 HPW/RHW

PHYSIOCOGNITIVE MODELING

The graphic above depicts several compounds that can and cannot be controlled by the warfighter. Measuring these compounds help predict the cognitive performance improvements to ensure maximum performance.

The Warfighters' cognitive performance on tasks can be hindered through exposure to detrimental chemical compounds, the restriction of critical compounds (e.g., oxygen), or enhanced through pharmaceuticals or nutraceuticals (e.g., caffeine mitigating the detrimental effects of fatigue). The goal of the *physiocognitive modeling* task is to derive cognitive performance predictions through the integration of computational models of physiology with computational cognitive process models. Some environmental compounds can be controlled by the Warfighter (e.g., caffeine intake during long-duration missions) and some cannot (hypoxic conditions during flight and toxin exposure during aircraft maintenance). The ability to predict the onset, magnitude, and duration of cognitive performance improvements—given different stimulant ingestion times and quantities—would facilitate maximum performance for long-haul flights or late-night cyber operations. Further, accurate cognitive performance predictions given increasing, decreasing, or stable oxygen or toxin levels would provide critical information to the Warfighter on whether to continue or abandon the mission based on safety and effectiveness considerations.

The process applied to achieve model-based predictions of compounds' effects on cognitive performance first requires computing simulated-blood concentrations of a targeted compound (oxygen, toluene, caffeine, etc.) using *physiologically-based pharmacokinetic* (PBPK) models. Next, the blood concentrations are mapped to computational cognitive process model parameters associated with specific cognitive capacities (e.g., attention, memory, executive control, motor control, etc.). This mapping is the major focus of the empirical and modeling research. Once a blood concentration-to-parameter mapping has been identified, cognitive performance across a set of different cognitive tasks that target the different capacities can be computed to determine which capacities show performance changes. The summary of performance changes across capacities can then be used to inform the Warfighter's commanding officer (e.g., readiness metrics) or the Warfighter directly (e.g., alarms and warnings).

To date, we have developed models of the effects of a common solvent on cognitive processing (toluene; Fisher, et al., 2017) and how caffeine impacts fatigue (Halverson, et al., 2021). Further, we have developed a cognitive task battery that targets different cognitive capacities for deployment in exposure studies. Data collected from these studies will facilitate model development by determining if, and to what extent, cognitive capacities are affected differently by exposure to different compounds. Currently, we're investigating approaches to modeling and predicting the effects of hypoxia on cognitive performance with partners from the University of Florida and the Mayo Clinic. We're also collaborating with the 711th Human Performance Wing's Airman Biosciences Division to collect data from Airmen and Guardians exposed to isopropanol in isolation, and we plan to collect data on how cognitive capacities are affected by G-Force. ☆

Dr. Christopher Myers, Senior Cognitive Scientist, 711 HPW/RHWM

THE GAMING RESEARCH INTEGRATION FOR LEARNING LABORATORY®

The Gaming Research Integration for Learning Laboratory (GRILL®)—under the Air Force Research Laboratory (AFRL), 711th Human Performance Wing, Airman Systems Directorate, and Warfighter Interactions and Readiness Division—leverages low-cost, high-fidelity, commercial off-the-shelf technology to create a variety of solutions for partners in the Department of Defense (DoD).

GRILL® OBJECTIVES:

Exploit existing commercial off-the-shelf & government off-the-shelf gaming tech to close the gaps between USAF and USSF training and simulations.

Identify emerging gaming tech and potential use cases.

Assist DoD partners with rapid proof of concept prototyping.

Mentor students and educators in science, technology, engineering, and mathematics (STEM) disciplines and develop workforce proficiency within the local area.

By pursuing these objectives, the team provides a greater return on investment for the DoD. Through utilizing existing technology to close training and simulation gaps, we reduce manpower costs by removing the need to develop tools from the ground up. Additionally, by utilizing platforms revolving around well-established game engines, the team ensures the continuity of our simulators by having tools that are frequently updated and compatible with emerging state-of-the-art hardware. The team's vision is to serve partners in the industry, academia, and the DoD as a resource for the integration and application of game-based hardware and software to address Air and Space Force needs. By way of collaboration with our partners throughout the United States, the GRILL® supports the development of training and simulation tools as well as highly-qualified technical and scientific workforce. ☆

Lt Kyle Bucklew, Program Manager, RHWM

Mr. Jon Diemunsch, Software Engineer, RHWM

Mr. Quintin Oliver, Development Engineer, RHWM

Dr. Winston "Wink" Bennett, Product Line Lead and GRILL® Team Lead, RHW

STUDENT SUMMER SHOWCASE

Over the summer, the GRILL® provided the opportunity for hands-on learning to students in the science, technology, engineering, and mathematics (STEM) fields. Through this program, the students used their skillsets to develop interactive programs with practical and effective applications towards situational-personnel training.

As a means to display the programs and subsequent hard work of the GRILL's® students, their work was shown to the public through the annual Summer Showcase. While the country still recovers from the effects of COVID-19, the GRILL® was able to host an in-person showcase of presentations and hands-on program demonstrations to the public.

Below and to the right are displays of the projects that were showcased during the summer open house.

SPACEX DRAGON CAPSULE ASTRONAUT RECOVERY SIMULATION

To help train for SpaceX Capsule recovery missions, GRILL® students developed a simulation where players must clear a checklist of tasks when retrieving or rescuing astronauts from a re-entered space capsule.

RESULTS: The students produced a retrieve-and-rescue simulation where players must complete a thorough checklist of tasks to ensure the astronauts are safe (while at sea). This included steps like properly removing the capsule's latch, recognizing chemical hazards, applying first aid, removing seat restraints, and stabilizing the capsule.

Photos by Ms. Shania Horner
Screen Captures by 711 HPW/RHW

SATELLITE PROXIMITY OPS

GRILL® developers were tasked with producing a simulation that put users in the position of controlling satellites while simulating a space-like environment impacted by realistic gravity physics.

RESULTS: The students produced a game where users dock satellites into each other while utilizing limited points-of-view. To increase immersion and engagement, players earn points and highscores based on the time and efficiency of their docking.

AR MAINTENANCE TRAINER

GRILL® students set out to build an AR maintenance trainer that overlays 3D models to real-life objects. With Unity® and the HoloLens 2®, equipment can be scanned into the software, while the hardware tracks the users' hand movements and eyes when performing maintenance.

RESULTS: Students created a smooth program interface with manually-animated 3D models that break down the equipment by each component. The software walks the user through the maintenance process—based on what they are fixing—by consulting a living, instructional data sheet on a local WAMPServer.

UAV DIGITAL ENGINEERING

Developers constructed a simulation for pilots to train with search-and-rescue UAVs, but the software had to be designed around the firmware and hardware of the drone. The simulation had to accurately represent the capabilities of the drone and simulate an environment that pilots would practice in.

RESULTS: Students made a simulation with an AI target that maps and alters its path, along with a drone that can carry livestreaming equipment and sensors that thermally screen the environment. Operators can register the body heat of their target and the environment's AI animals, resembling what pilots will experience on a search-and-rescue mission with drones.

VR PILOT STRAINING

To measure pilot fatigue, developers produced a simulator that allows users to accurately measure their muscle tension. Ultimately, the developers wanted to make the simulation for pilots to learn how to brace during a flight while preventing the in-simulation character from blacking out due to G-LOC.

RESULTS: Using a 360-view video in VR, the developers made a SIM equipped with metrics that measure muscle tension and grip strength through a bio-radio held within the right hand. The according arm is given three electric medical pads that provide accurate, on-screen readings of body tension.

PERSONALIZED LEARNING & READINESS SCIENCES (PLRS)

Ms. Jennifer Winner
Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL

TWO PROJECT AREAS

- **Adaptive Proficiency Technologies**
- **Interactive Task Learning**

Both projects involve multiple lines of effort, each with its unique and complementary role to play in creating new capabilities for the future fight.

The PLRS ("pillars") CRA is focused both on scientific discoveries related to the fundamental nature of readiness and the exploratory application of personalization technologies for more effective and efficient learning, improved knowledge and skill retention, and more robust mission readiness. Our focus is on enabling improved learning and performance for robust mission readiness for human and human-machine teams. The technical challenges we are tackling within PLRS are organized around two themes:

1) Adaptive Proficiency Technologies (APT)

2) Human and Machine Co-Learning (Co-learning).

Top priorities within APT research include the definition and objective quantification of readiness standards for individuals and teams, optimization across curriculum selection and timing to maximize improvement, retention and relearning, and the implementation of personalized learning in USAF and USSF use cases. Regarding Co-learning research, our top priorities include defining new paradigms for integrating emerging AI technologies into intelligent tutoring, human-machine co-training and interactive learning, AI-enabled apprenticeships, and establishing techniques for machine goal-inference from humans in a common mission environment. Within both of these lines of effort, digital training data are key enablers. The availability of digital data is critical for achieving high-resolution human and system measurements and proficiency-centric readiness evaluations for individuals and unit-level assessments and feedback. The following articles illustrate specific examples of our current investment in these areas. ☆

Ms. Jennifer Winner, Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL

PLRS OVERVIEW

ADAPTIVE PROFICIENCY TECHNOLOGIES

In preparation for peer fights, during which just-in-time training requirements for novel tasks are likely to be the norm, methods for training and ensuring adaptive performance are critical."

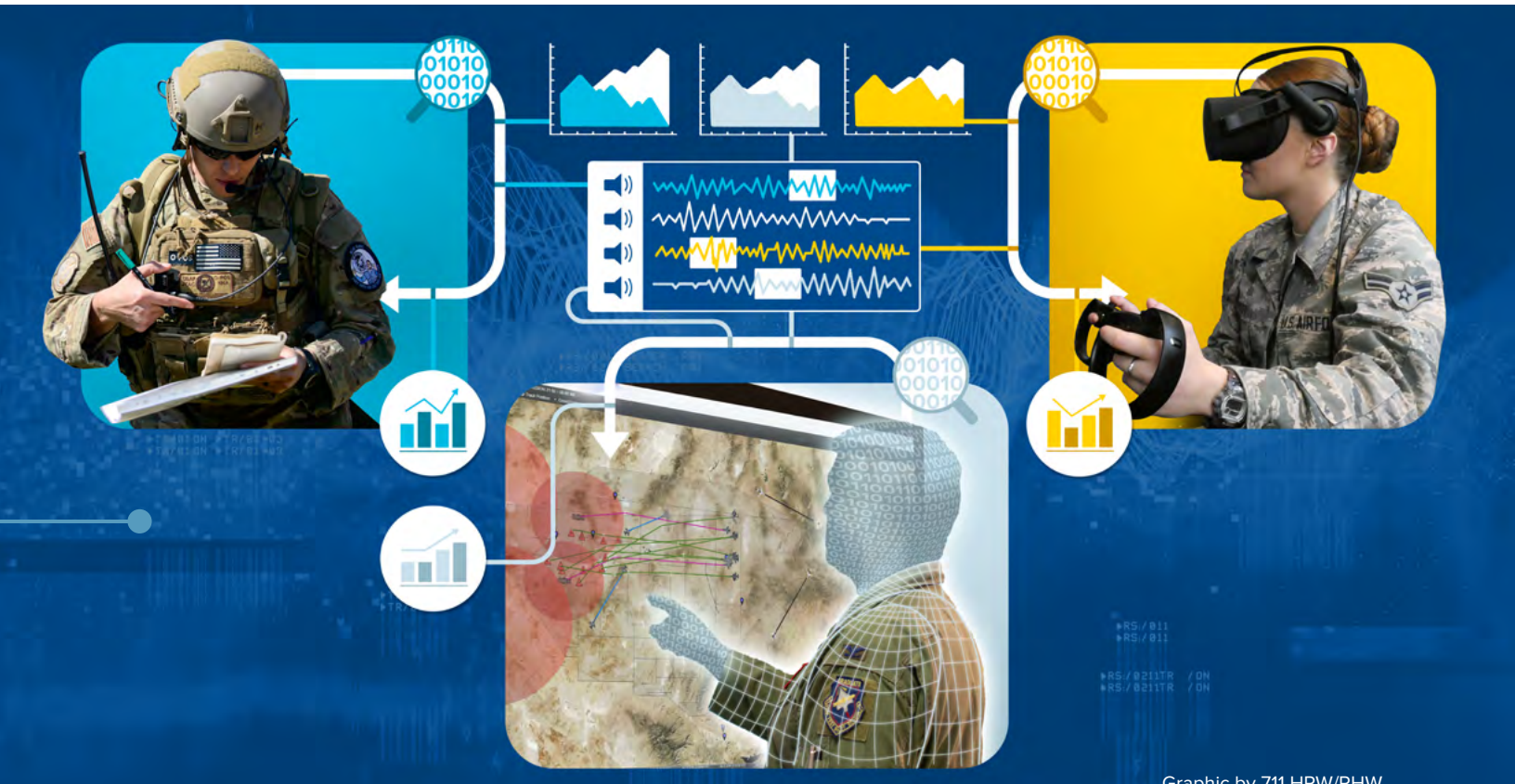
—Ms. Jennifer Winner
Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL

ERROR RECOGNITION AND RECOVERY TRAINING

A top priority within the Adaptive Proficiency Technologies line of effort is better tools and recommendations related to readiness standards for individuals. In preparation for peer fights, during which just-in-time training requirements for novel tasks are likely to be the norm, methods for training and ensuring adaptive performance are critical. Our team has partnered with Applied Decision Sciences to begin to identify and investigate training methods and scenario design approaches with potential to maximize performance transfers to novel tasks. Ms. Laura Militello and team have completed an analysis of literature that summarize scenario design options related to error recognition and recovery training, or error management training as it is referred to in the medical training literature. The findings of this literature will be tested in an empirical study looking at the efficacy of these training approaches and their promise for adaptive transfer (i.e. the transfer of performance on novel tasks). If findings from the literature are upheld in the empirical tests, these scenario design seeds may be leveraged to support and maximize transfer of training for novel tasks in complex mission types that require just-in-time training. ☆

Ms. Jennifer Winner, Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL

Ms. Laura Militello, Senior Scientist Contractor, 711 HPW/RHWL



Graphic by 711 HPW/RHW

PERFORMANCE-ENABLED OPERATIONAL TRAINING ENVIRONMENT (PEOTE)

PEOTE's task is to research, engineer, and implement personalized-learning capabilities in USAF & USSF use cases by defining best practices for assessing and implementing advanced-learning technologies (e.g., augmented reality (AR) and mixed reality (MR)) and enabling digital training data to achieve high-resolution human and system measurements, along with proficiency-centric readiness evaluations and feedback. A key part of this task is exploring ways to capture, classify, and categorize objective data to populate data sets across operational and tactical mission areas.

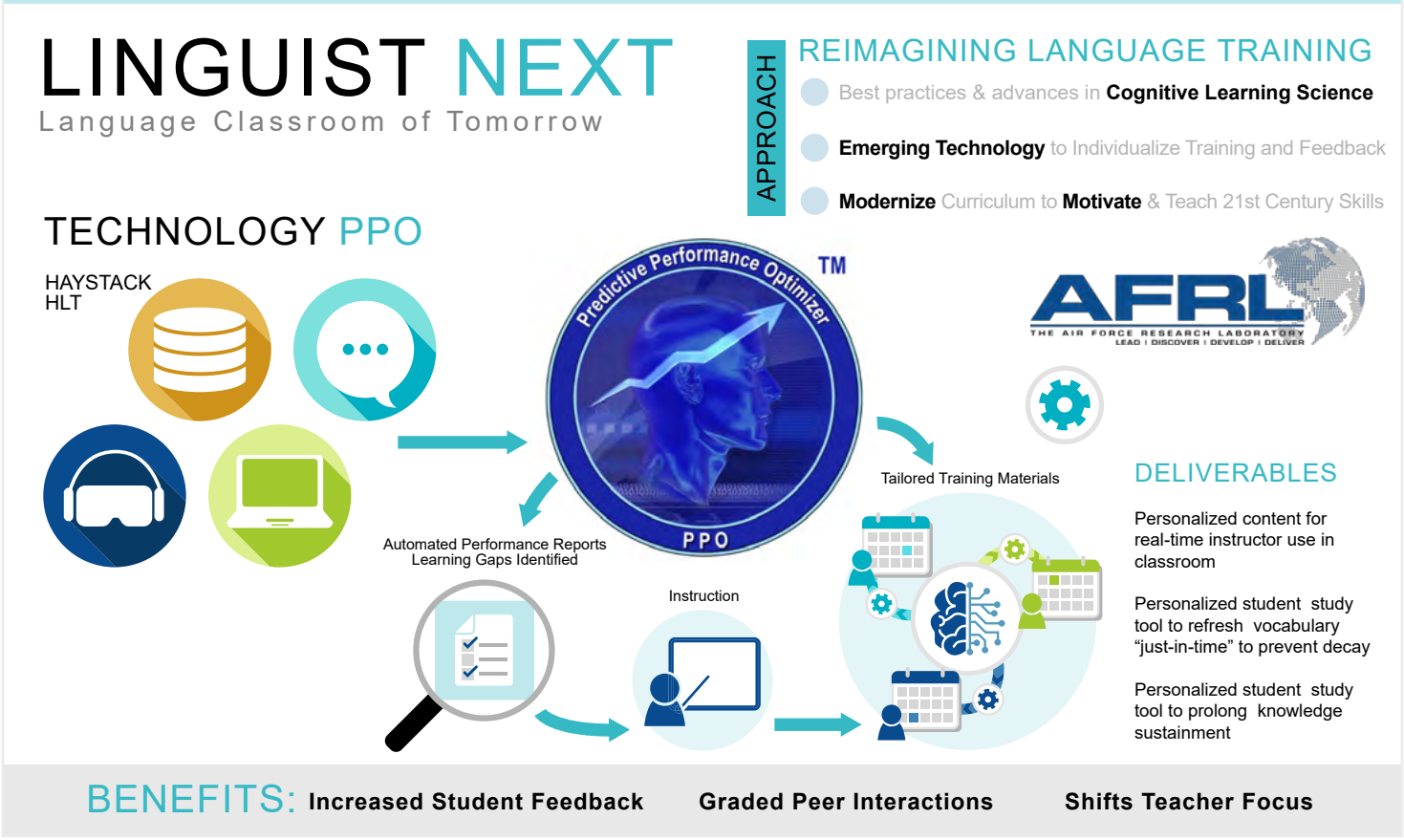
PEOTE experiments with several different LVC technologies while testing training data collection methods in operationally representative environments. Our experimentation includes live and synthetic effect employments and capturing interactions, situational awareness, and readiness assessments of the training participants in multi-domain (space, cyber, air, and ground) environments. Also, as part of its research mission, PEOTE partners with the Gaming Research Integration for Learning Laboratory (GRILL®) to apply emerging AR and VR technologies in live scenarios to determine the limits of these technologies, the boundary conditions that limit their effectiveness, and to frame future requirements for uses in training.

In the near term, the PEOTE team is building an environment-mapping capability that will merge live and synthetic spaces. Enabling the real-time synchronization of the live and synthetic environments will ensure greater fidelity and inclusion of mission partners, regardless of their participation status (live, simulator, cyber, etc.) Our dual-use data collection and LVC training effects platform will have autonomous environmental-mapping and entity-classification capabilities. Our developmental-testing phase will take place at our Calamityville site prior to deployment at more complex operational training locations. This capability will enable the PEOTE team to collect vast amounts of data from the training environment; hone and convert that data to fundamental components; and then use those components to characterize and synchronize the live and synthetic environments while simultaneously allowing the exploration of objective performance measures at abstracted mission levels. ☆

Mr. Ted Harmer, Electronics Engineer, 711 HPW/RHWL

The Calamityville site acts as a live test environment that we can map and characterize, convert into a virtual representation where live participants can interact, and create effects that are then translated back into the live environment.

Graphic by Mr. Will Graver



Graphic by Dr. Julie Cantwell

OPTIMIZED LEARNING FOR LINGUISTS

Research investment under our Adaptive Proficiency Technologies line of effort within PLRS includes focus on the optimization of learning, retention, and relearning through adaptation of curriculum selection and timing. Central to this work are maturation efforts focused on the Predictive Performance Optimizer (PPO), which is a [state-of-the-art cognitive model designed to track learning and decay in order to make personalized and tailored study recommendations to individual learners]. PPO, patented in 2013, has been demonstrated to enhance performance effectiveness and minimize training time through field study applications with the American Heart Association. Recent maturation efforts are evaluating utility of the approach across key DoD-specific use cases, including linguist training and total-force resiliency training.

Predictive Analytics for Learning (PAL) researchers are piloting an innovative integration of the PPO's personalized learning capabilities with the automated translation capability of RHWAI's Haystack Technologies' foreign media as part of the Air Education Training Command's (AETC) Linguist Next project. Arabic students at the Defense Language Institute (DLI) will be leveraging this technology over the course of their 64-week training curriculum.

Personalization in the classroom will unfold in two specific ways. First, the technology will provide instructors with real-time performance assessments and (based on PPO-identified student weaknesses) will point to Haystack to provide optimized instructional material recommendations to remediate known deficiencies. Second, students will leverage a novel app outside of the classroom to optimize the study of both new and review vocabulary items. This is done by automating the acquisition and sustainment of specific word pairs based on student performance history and prioritization of content that's based on instructor-identified needs, statistical utility, or syntactic/semantic relations to previously learned items.

Our team is diligently working with DLI and Linguist Next experts to digitize all learning and assessment materials in order to track student performance at a fine grain of analysis and determine whether the integrated technology creates higher-performing linguists compared to a control group of DLI students. We hope that if positive evidence is attained, our team may transition this technology to DLI at large and expand to the full suite of target languages. ☆

Dr. Tiffany Jastrzembski, Team Lead and Senior Cognitive Scientist, RHWS

Dr. Julie Cantwell, AETC Linguist Next Project Director

Dr. Florian Sense, Cognitive & Computer Scientist, Infinite Tactics

Mr. Michael Krusmark, Research Scientist, CAE USA

Dr. Jong Kim, Research Scientist, CAE USA

Mr. Michael Collins, Research Scientist, Oak Ride Institute for Science and Education

TEAM DYNAMICS MEASUREMENT

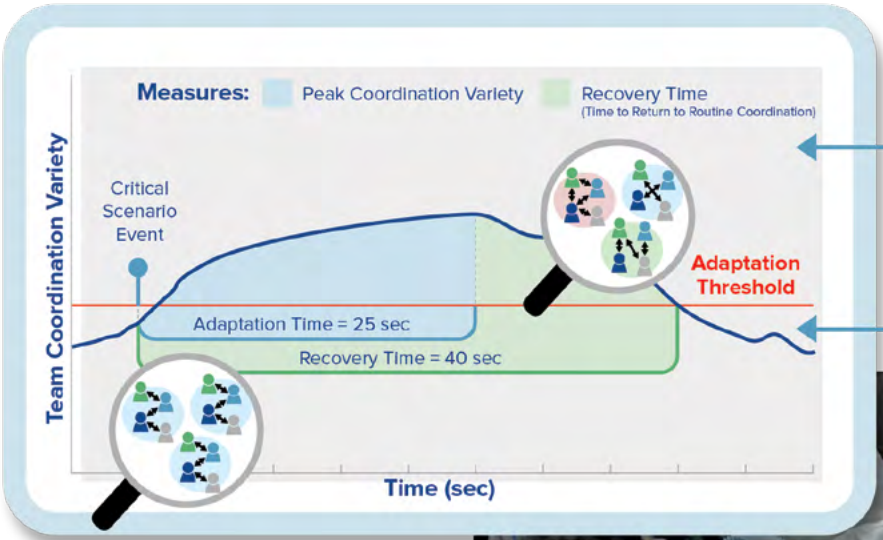
One of the top priorities within the Adaptive Proficiency Technologies line of effort is to define and objectively quantify readiness standards for individuals and teams. In real-world training environments, there is a lack of methods that disentangle individual contributions toward team coordination performance in complex mission scenarios. Our team has partnered with the Georgia Institute of Technology to apply existing measures of team coordination performance in USAF use cases. The first use case in which we are exploring this capability is Critical Care Air Transport (CCAT) teams. In partnership with the USAF School of Aerospace Medicine En Route Care Training Department (USAFSAM/ET), we will be demonstrating quantitative measures for team coordination behavior and developing a real-time assessment tool to supplement the information available to instructors. The data will support the deliberate practice and feedback model that defines the CCAT courses. Future use cases for this capability will build upon the success of this tool in providing sufficient discrimination between levels of achievement for team coordination. ☆

Ms. Jennifer Winner, Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL

Dr. Jayde King, Principal Investigator and Research Psychologist, 711 HPW/RHWL

Dr. Jamie Gorman, Associate Professor of Engineering Psychology, Georgia Institute of Technology

> **Future use cases for this capability will build upon the success of this tool in providing sufficient discrimination between levels of achievement for team coordination."**
>
> **— Ms. Jennifer Winner**
> Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL



ADAPTATION THRESHOLD

Team Coordination Above the Adaptation Threshold Marked by Increased Coordination Pattern Variety

Team Coordination Below the Adaptation Threshold Marked by Stable and Consistent Coordination Patterns

Graphic by Ms. Shania Horner

INTERACTIVE TASK LEARNING

MISSION PLANNING AND DEBRIEF

The Mission Planning and Debrief team is leading research, development, and evaluation efforts for next-generation mission planning and debrief for the near peer-contested fight. The goal is to streamline mission planning processes and promote human-AI collaboration to improve mission outcomes. Our approach involves combinations of new technologies that allow human-AI teams to engage in these iterative processes together. We named our prototype application "Metis."

Metis provides a variety of communication methodologies, with a few examples being a web-based interface, RESTful API, and Websockets. Our communication interfaces enable human, external system, and AI integration and are connected to the same backend that synchronizes all incoming data and immediately propagates information to all connected endpoints in real-time. These capabilities have been developed in a service-oriented architecture, and our team is developing an array of machine intelligence and automation services, including the following:

- Mission plan generation agent
- Mission plan validation agent
- Constrain extraction agent
- Planning product development service

These in-development technologies are at relatively early stages of maturity, but we have already begun formative evaluations. The evaluations indicate dramatic improvements in workflow and subsequent decreases in process completion time. Early comparisons of machine-generated content to human-generated content show that they are comparable.

Photo by Mr. Will Graver

As part of our ongoing core applied research, the Mission Planning and Debrief laboratory space for evaluation, along with the current implementation of associated agent technologies and services, provides a foundation to create the future for human-AI collaboration. From the perspective of planning the research program, the concept is to use the lab space as a research and development platform as well as leverage its infrastructure for new investments. These investments span the continuum from fundamental research to advanced technology demonstrations as time, attention, and funding allow.
Preliminary results are encouraging, and we are enthusiastic about the possibility of near-term technology off-ramps that can benefit training at large-force exercises. However, the end goal is to provide technology options that benefit operations and provide flexible force readiness more broadly. Our aim is to move the technologies toward multi- and all-domain operations and a broader set of functional mission types and associated assets. ☆

Mr. Sean Kennedy, Computer Scientist, 711 HPW/RHWL

Graphic by Ms. Shania Horner

CARNEGIE MELLON UNIVERSITY OPERATING LOCATION

Meeting tomorrow's technological challenges for the Warfighter requires laying a solid foundation of relevant, innovative basic research today. The Warfighter Readiness & Interactions Division is thinking outside the fence by embedding a researcher on the campus of Carnegie Mellon University to directly collaborate with technical thought leaders and leading-edge laboratories in cognitive modeling, artificial intelligence (AI) and human-autonomy teaming. Our experiment in changing the way we collaborate with universities is manifesting the goals of expanding the USAF's presence in innovation hotspots and enhancing our talent recruitment, as outlined in the *USAF Science and Technology Strategy 2030.*

The Carnegie Mellon University (CMU) Operating Location has been operational for three years, since fall of 2018. In that time, we have initiated several collaborations, made possible by a Cooperative Research and Development Agreement in the area of Cognitive Modeling for Human-Autonomy Teaming. This has facilitated new research in statistical methods for robust cognitive modeling and recurrence quantification-based characterizations of decision-making strategies with researchers in the Department of Psychology and the School of Social and Decision Sciences. We've started to investigate the ways robustness can be measured for modeling and simulation with researchers in the Institute for Software Research. Also, we've developed a framework for modeling the calibration of trust in autonomy that also enables assessments of the vulnerabilities the calibration process creates for adversarial exploitation with the Department of Psychology.

Carnegie Mellon University is also home to two AFRL Centers of Excellence (COEs), including the 711th HPW-supported Center of Excellence on Trusted Human-Autonomy Teaming. Our presence on campus enables direct collaborations with the COEs, facilitating even more creative advances to our knowledge of how to use artificial intelligence to enhance our human knowledge, skills, and readiness.

In addition to building awareness of CMU research, our presence at the CMU Operating Location facilitates engagements with the growing Pittsburgh tech community that's heavily invested in robotics. We can engage with DARPA-challenge winning teams and the US Army AI Task Force at the National Robotics Engineering Center and see emerging technologies in self-driving cars that are out for test drives in the local neighborhoods. CMU is an inspiring place to be!

Individuals interested in connecting to the Center of Excellence on Trusted Human-Autonomy Teaming are invited to join the monthly AFRL-COE Tech Talks, held the first Thursday of each month at 1200 Eastern on Zoom®. See below for contact information! ☆

Dr. Leslie Blaha, Senior Research Psychologist and CMU Operating Location Lead, RHWL and Carnegie Mellon University

TECH TALK CONTACT INFORMATION:

leslie.blaha@us.af.mil

FIDELITY EVALUATION METHOD FOR SIMULATOR ASSESSMENT

Choosing the best simulator to train complex procedures is challenging. Cost, data accessibility, and procedural flexibility are among the many factors that must be considered to foster efficient, personally adaptable learning. Ideally, the selection of task simulators for use during training should be grounded in science that demonstrates realistic, procedural cue presentation when compared to the context of real-world performance.

The PLRS Team adopted and refined a simulator-fidelity assessment method (drawn from legacy AFRL research recommendations) that combines expert-informed task analysis with experimental data collection. The method associated the variations of sensory-cue presentation with requirements for skill mastery, along with the transfer of performance to applied situations. Researchers constructed instruments for task analysis and data collection that were sensitive to variations in the procedure's completion speed, errors of commission and omission, and eye-movement recording. These measures are ostensibly attributable to interactions with simulator features. During a transfer of training experiment with experienced and inexperienced task performers, participants completed a 90-minute experimental training and testing regimen. As part of the protocol, they supplemented their performance data with detailed insights about the simulators that will benefit future trainer refinements and guide training designers and procurement specialists.

Analyses of completed self-reported and performance data from the initial-use case have led to several conclusions. Key among these was that capturing and analyzing data from a variety of sources facilitated meaningful simulator comparisons during training and testing. Applying advanced statistical methods made it possible to associate existing and missing simulator features to capabilities and limitations related to trainee instruction.

Since the fidelity evaluation method is grounded in sensory cue presentation, successful applications of the method are expected across various task domains and trainer types. Subsequent investigations are planned to target simulators that harness augmented and virtual reality. Research team members are confident their findings will enhance readiness assessments, providing valuable guidance for members of the military training and procurement communities as they strive to ensure rapid and comprehensive acquisition of critical skills by trainees. ☆

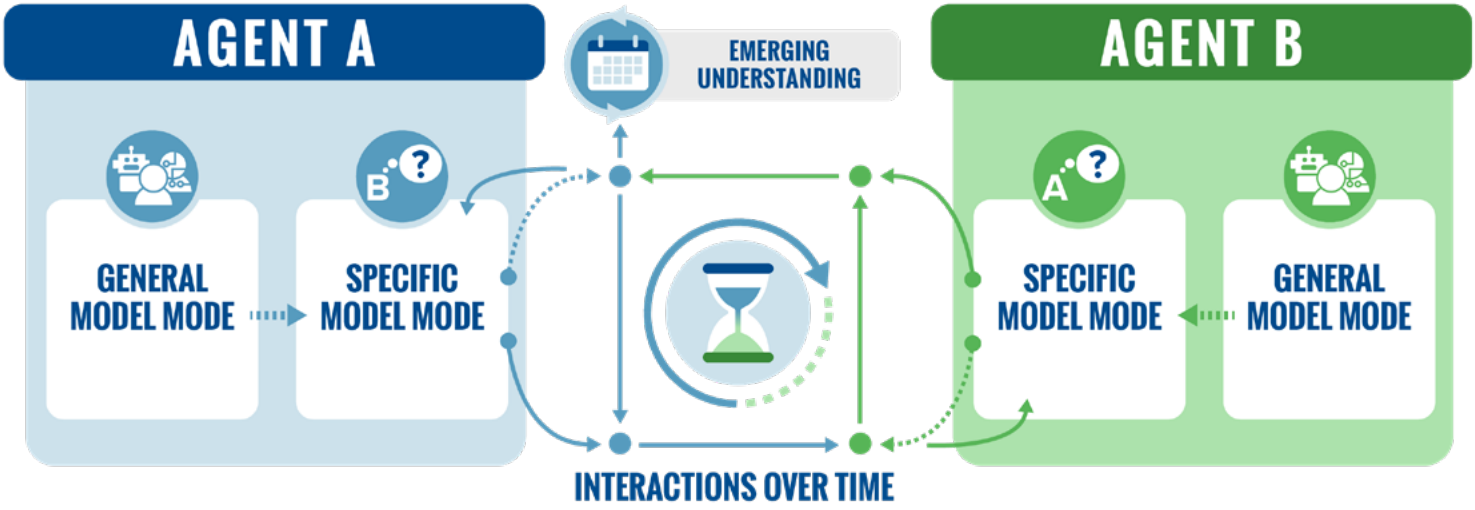
Ms. Jennifer Winner, Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL

Dr. Jim Bliss, TO-19 Lead, Leidos

Mr. Kent Etherton, Research Scientist, Leidos

Photo by Mr. Will Graver

> **Research team members are confident their findings will enhance readiness assessments, providing valuable guidance for members of the military training and procurement communities as they strive to ensure rapid and comprehensive acquisition of critical skills by trainees."**
>
> **— Ms. Jennifer Winner**
> Core Research Area Lead for Personalized Learning and Readiness Sciences and Principal Cognitive Scientist, 711 HPW/RHWL



Graphic by Ms. Shania Horner

THE SCIENCE OF UNDERSTANDING

Future Airmen and Guardians will train and execute their missions in collaboration with intelligent machines. The vision for seamless, effective teaming and co-learning assumes that both humans and machines understand each other fluidly to predict each other's decisions, actions, and future behaviors. However, what does it mean for machines and humans to understand each other and how do we measure understanding successes and failures? Since understanding is a multi-faceted, and sometimes ill-defined concept, we do not yet have systematic methods to assess if we have achieved desired levels of understanding in humans, intelligent machines, and human-autonomy teams.

The research within The Science of Understanding program is laying a rigorous foundation for the assessment of understanding. We are exploring the multidisciplinary, conceptual space-encompassing varieties of understanding by identifying the similarities and differences between human and machine "mental models," and developing a new Theory of Mutual Understanding. We seek to define generalizable tests and metrics to measure understanding, both in laboratory and under mission-relevant teaming tasks. We are advancing statistical techniques for quantifying behaviors over time and quantifying the coordination of language, motor, and decision-making behaviors between agents.

The ways humans and intelligent machines gather information and reason about the world and other agents are not the same. We are pursuing the Science of Understanding with special attention to the fundamental similarities and asymmetries among humans and machines understanding each other. We adopt the perspective that cognitive models have a critical role to play in facilitating machine understanding of humans. In several ongoing efforts, we are demonstrating how models instantiated in computational cognitive architectures help to characterize human decision-making strategies, predict how diversity of learning can lead to better skill generalization, and quantify human perceptions of an autonomous system's reliability and explore the vulnerabilities of that calibration process (see publications).

The Science of Understanding is a fundamental research collaboration between scientists in the Warfighter Interactions and Readiness Division, Naval Research Laboratory, Army Research Laboratory, Carnegie Mellon University, and the University of Minho; it is the flagship research program for AFRL's Carnegie Mellon University Operating Location. ☆

TEAM MEMBERS:

Dr. Leslie Blaha, Senior Research Psychologist and CMU Operating Location Lead, RHWL and Carnegie Mellon University

Dr. Sarah Bibyk, Research Linguist, RHWM

Capt Mitchell Cochell, Program Manager, RHWL

Mr. Patrick Dull, AI/Simulation Software Engineer, CAE

Dr. Coty Gonzalez, Research Professor, Carnegie Mellon University

Dr. Beth Hartzler, Research Psychologist, CAE

Dr. Jayde King, Human Factors Engineer, RHWL

Dr. Christian Lebiere, Research Professor, Carnegie Mellon University

Dr. Erin McCormick, Research Fellow, Carnegie Mellon University

Dr. Alfredo Pereira, Research Scientist, University of Minho

Ms. Rachel Wong, Research Scientist, CAE

RECENT PUBLICATIONS

Blaha, L. M., Lebiere, C., Fallon, C. K., & Jefferson, B. (2020). Cognitive Mechanisms for Calibrating Trust and Reliance on Automation. Proceedings of the 18th International Conference on Cognitive Modeling

Lebiere, C., Blaha, L. M., Fallon, C. & Jefferson, B. A. (2021). Adaptive cognitive mechanisms to maintain calibrated trust and reliance in automation. Frontiers in Robotics, 8, 135. https://doi.org/10.3389/frobt.2021.652776

McCormick, E., Blaha, L. M., & Gonzalez, C. (2020). Exploring Dynamic Decision Making Strategies With Recurrence Quantification Analysis. Proceedings of the 42nd Annual Meeting of the Cognitive Science Society

McDonald, C., Gonzalez, C., Blaha, L.M., Lebiere, C., Fiechter, J., Bugbee, E., & McCormick, E. N. (2021). Diverse experience leads to improved adaptation: An experiment with a cognitive model of learning. Proceedings of the 19th International Conference on Cognitive Modeling (ICCM 2021).

NO SEE, NO HEAR, NO TEAMMATE?

Enabling Goal Inference for the Autonomous Teammates of Tomorrow

Rigid, unadaptable, autonomous agents/systems have been a long-standing technical challenge within the human and autonomous teaming literature. Most autonomous agents today are unable to perceive and anticipate user needs, goals, and status within a task. This lack of perception can lead to poor awareness from autonomous agents about their human teammates, ultimately resulting in poor human-machine collaboration and performance. According to the *Air Force Science and Technology 2030 Strategy,* artificial intelligence and human machine teaming are essential to equipping the future Warfighter with rapid decision-making capabilities. In response to the need to accelerate change, the Human and Machine Co-Learning line of effort is researching techniques for improving collaboration between humans and autonomous agents and discovering new ways to promote simultaneous human-machine learning.

Our Vision Inference Action (VIA) research is developing an ensemble model that leverages visual Bayesian inferencing and natural language processing techniques to provide a more robust goal-inferencing capability for systems and agents. This summer, we made considerable progress on both the visual Bayesian inference model and the natural language processing model. Now, we are exploring methods to combine these two models while maintaining system transparency and cohesion. We are also investigating effective approaches for using the information gained from our ensemble model to equip autonomous agents to inform, instruct, correct, and alert users. Future research will include identifying mission-relevant use cases for these tools and developing a synthetic task environment to support human and autonomous agent research. ☆

TEAM MEMBERS

Dr. Jayde King, Principal Investigator and Research Psychologist, 711 HPW/RHWL

Captain Mitchell Cochell, Program Manager, 711 HPW/RHWL

Mr. Sean Kennedy, Computer Scientist, 711 HPW/RHWL

Dr. Leslie Blaha, Senior Research Psychologist and CMU Operating Location Lead, RHWL and Carnegie Mellon University

CONTRACTORS

Mr. Jimmy Cline, Software Engineer, CAE USA

Mr. Alex Hough, Cognitive Scientist, ORISE and Wright State University

Mr. Michael Collins, Cognitive Scientist, ORISE and Wright State University

Dr. Michael Lee, Cognitive Scientist, ORISE and UC Irvine

Ms. Lauren Montgomery, Cognitive Scientist

Ms. Helen Yang, Cognitive Scientist

Graphic by 711 HPW/RHW

COMMANDER'S RESEARCH AND DEVELOPMENT FUNDS EFFORT

Commander's Research and Development Funds (CRDF) are enterprise-level investments in topics of high priority that are reviewed by the Research Advisory Council and approved by the AFRL Commander.

ROBUST AND SECURE MACHINE LEARNING

The Air and Space Forces are leaning into emerging machine learning (ML) and artificial intelligence (AI) technologies and applications to improve Warfighter capabilities in many domains. While recent breakthroughs present great opportunities, they also introduce new challenges. One challenge is that the fragilities and vulnerabilities of deep-learning technologies are not yet well understood; emerging research has shown a less-thoughtfully designed and tested ML model can be easily exploited for significant performance degradation. Another challenge is quantifying the robustness of ML in the face of such exploitations and identifying mechanisms that improve the robustness of ML systems. An additional challenge is developing a thorough understanding of the ways human teammates might improve the robustness of ML-enabled systems and how ML-enabled system exploitation may impact human-AI team operations. Therefore, as identified in the *2018 DoD Artificial Intelligence Strategy*, it is critical for the Air and Space Forces to invest in "resilient, robust, reliable, and secure" AI/ML technologies to maximize performance robustness, analyze and minimize security vulnerabilities, and serve as counter-measures against adversarial systems.

Researchers from the Airman Systems Directorate, led by RHW, have partnered with researchers from the Information, Sensors, and Munitions Directorates in an AFRL-wide effort from the Commander's Research and Development Funds to invest in robust and secure ML. The R&D commenced in 2020. We adopted a teaming approach that divides responsibilities across Red Team offensive attacks, Blue Team defensive models, and Green Team objective evaluation. Given our history of research of conceptualization and quantification of robustness, along with novel approaches to interactive ML, our Airman Systems Directorate team serves predominantly in the Green Team role.

During the first year of our efforts, we established a cross-directorate computational infrastructure supporting the first annual all-team challenge event. This event pitted the early Red and Blue team technologies in a series of image classification challenges that assessed robustness to adversarial ML attacks drawn from cutting edge ML literature and to additional sources of noise system perturbation. We successfully deployed multi-dimensional assessments of robustness, including the first test of a novel metric, the Index of Robustness. This challenge event firmly established a baseline for existing ML capabilities and identified critical directions for the remaining research program.

In 2021, our team continues to develop methods for conducting multi-dimensional assessments of offensive and defensive ML robustness. In order to quantify and characterize ML robustness, we are pushing the state-of-the-art in methods for quantifying human and machine perception and classification performance. We also delved into evaluating robustness pros and cons of human-in-the-loop active- and interactive-ML training and classifiers. We are developing novel interactive-ML interfaces to leverage in future challenge events. Progress in 2021 will (again) culminate in an annual, all-team demonstration event. ☆

THE ROBUST AND SECURE MACHINE LEARNING RHW TEAM:

GOVERNMENT:

Dr. Leslie Blaha, Senior Research Psychologist and Carnegie Mellon University Operating Location Lead, 711 HPW/RHWL

Mr. Michael Dougherty, Computer Scientist, 711 HPW/RHWL

Mr. Sean Kennedy, Computer Scientist, 711 HPW/RHWL

Dr. Jayde King, Research Psychologist, 711 HPW/RHWL

IN-HOUSE CONTRACT:

Mr. Patrick Dull, AI / Simulation Software Engineer, CAE

Dr. Marla Kennedy, AI Researcher, Leidos

Dr. Platon Lukyanenko, AI Researcher, CAE

Mr. Ryan McCoppin, Senior Data Scientist, CAE

Ms. Meghan Sorensen, Software Engineer and TO12 Lead, CAE

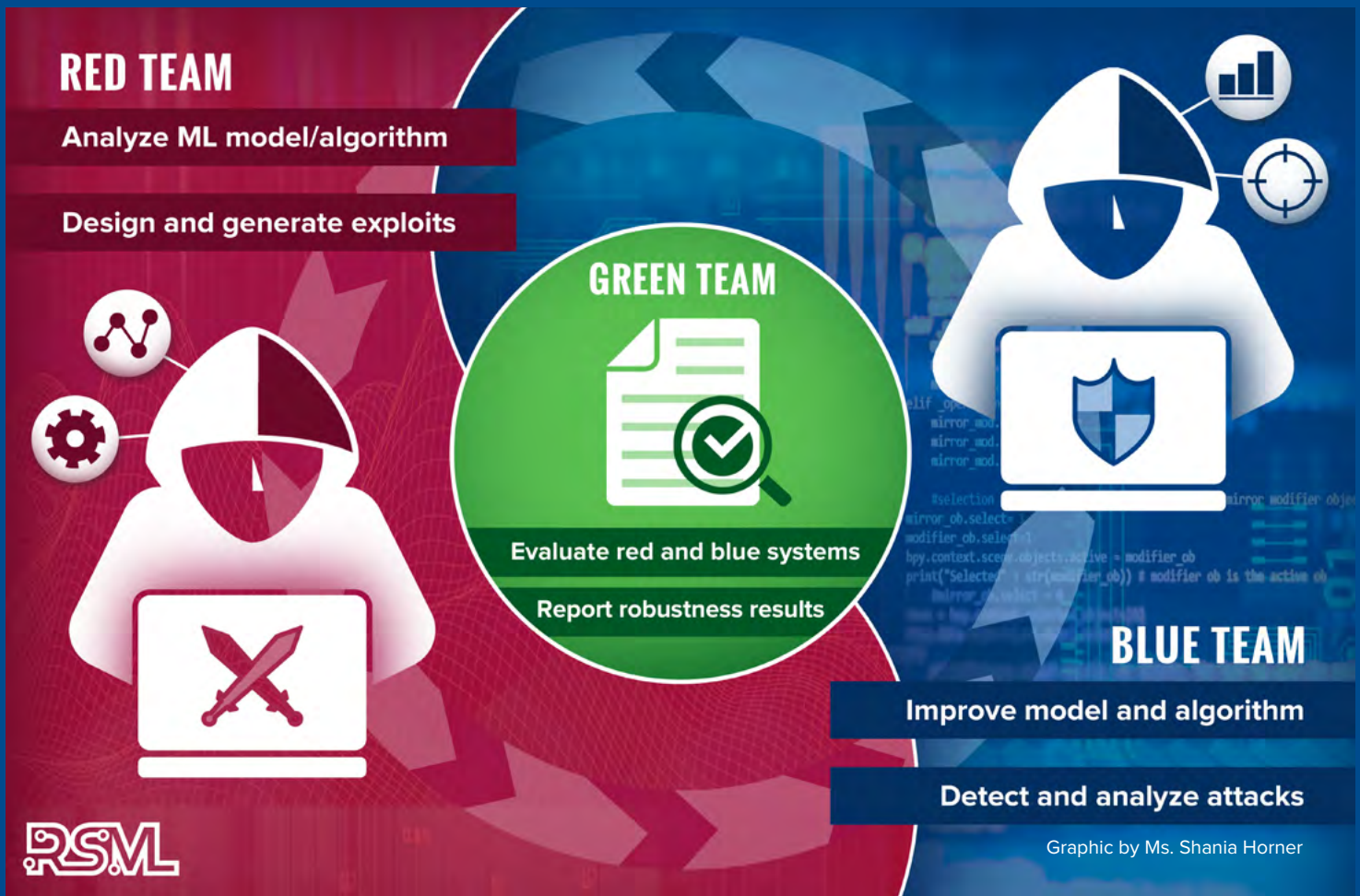
INTERNS:

Mr. Alexander Dimopoulos, Repperger Research Intern, University of California San Diego

Mr. Jacob Gruver, ATR Center Intern, Boston University

[THE ALL-TEAM] CHALLENGE EVENT FIRMLY ESTABLISHED A BASELINE FOR EXISTING [MACHINE LEARNING] CAPABILITIES AND IDENTIFIED CRITICAL DIRECTIONS FOR THE REMAINING RESEARCH PROGRAM."

— Dr. Leslie Blaha
Senior Research Psychologist and
Carnegie Mellon University Operating Location Lead,
711 HPW/RHWL



Graphic by Ms. Shania Horner

ADAPTIVE WARFIGHTER INTERFACES

OUR CTC EXISTS TO MAINTAIN DECISION SUPERIORITY THROUGH OPTIMAL WARFIGHTER INTERFACE DESIGN."

— Dr. Mark Draper
Adaptive Warfighter Interfaces Core Technical Competency Lead and Principle Engineering Research Psychologist, 711 HPW/RHW

Dr. Mark Draper
Adaptive Warfighter Interfaces Core Technical Competency Lead and Principle Engineering Research Psychologist, 711 HPW/RHW

Photo by AFRL

ADAPTIVE WARFIGHTER INTERFACES CTC

Operator interfaces are prevalent throughout the entirety of the Air and Space Forces; they provide the critical connections between warfighter and machine systems for all mission applications. Airmen and Guardians employ warfighting power by engaging and "interfacing" with increasingly complex and intelligent machines and information sources while under constant threats of data uncertainty and information subterfuge. Given that future wars will increasingly be wars of cognition, victory will likely be decided by those who best synergize human cognition and machine intelligence/capabilities to effectively operate within their opponent's decision cycles. Our CTC exists to maintain decision superiority through optimal Warfighter interface design.

The Adaptive Warfighter Interfaces (AWI) CTC discovers, demonstrates, and transitions the full-spectrum of operator interface science and technology, communication enhancements, and human-centered decision analytics that dynamically integrate Warfighters and increasingly intelligent machines into maximally-effective collaborative teams.

Automation and emerging AI technology can be exceptionally powerful tools when operating within known and well-defined situations/environments, but they frequently fail when operating outside of their competency envelopes. Sadly, in many cases, system designers follow a "leftover principle" where automation/AI is regularly applied and the human role is reduced to addressing "leftover" failures and gaps in automation functions. Thus an "out-of-the-loop" Warfighter is expected to identify intermittent automation failures and rapidly "jump in and save the day," with little time and little situation awareness of the problem at hand. This is exacerbated by automation/AI that's poorly designed with little-to-no transparency as to its functioning/reasoning. An alternate approach to system design is to consider a "joint cognitive system," where human cognition is more effectively and dynamically interleaved with complex system capabilities throughout the task space. The AWI CTC is dedicated to conducting R&D that directly informs these joint cognitive systems by focusing less on normative operations (where it is easy to succumb to the "leftover" principle) and concentrating on the highly complex, uncertain, and "edge-case" environments that will characterize future warfare.

THESE COMPLICATED ENVIRONMENTS WILL LIKELY INCLUDE:

Distributed operations with mixed, multi-echelon teams of humans and intelligent machines.

Fallible machines that continually gain intelligence, learn, and adapt over time.

High-stress environments requiring rapid, accurate situation assessment & decision making.

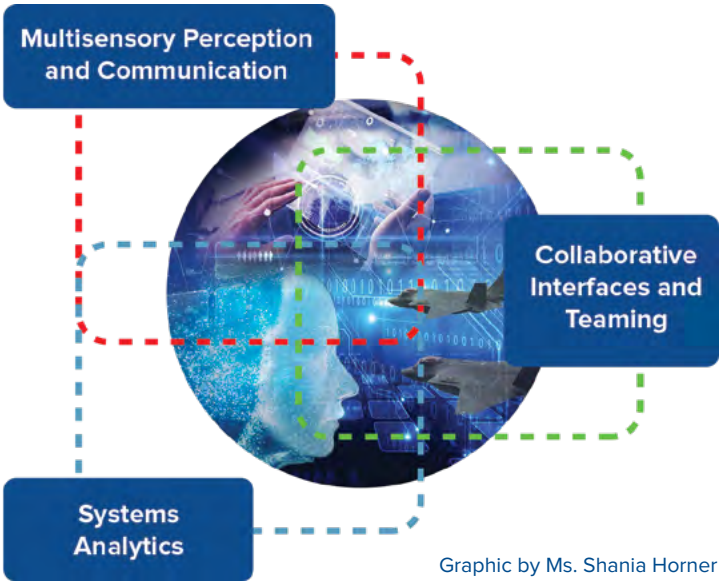
Situations marked by massive data overload that occludes relevant information.

Rapidly changing information & communication environments with varying uncertainty.

To successfully achieve its goals, AWI is organized into three, interrelated CRAs. The Multi-sensory Perception and Communication CRA focuses predominately on the "human side" of the system, with emphasis on characterizing and exploiting human perceptual abilities as well as enhancing communication capabilities within teams of humans and machines. The Systems Analytics CRA considers attributes from the "machine-side" of the system, emphasizing rapid "meaning-making" from large, complex datasets while

Continued on Next Page...

improving "analytic-enabled" cognition through more effective and user-centered design of machine analytics. Finally, the Collaborative Interfaces and Teaming CRA blends aspects of both humans and machines in the design of novel human-automation collaboration methods and distributed heterogeneous teaming interface solutions.



Graphic by Ms. Shania Horner

To enhance the value proposition of this CTC, AWI focuses on advancing and aggregating individual research efforts into larger, integrative solutions that enable high-priority Air & Space Force mission capabilities, such as joint all domain operations.

OUR RESEARCH PORTFOLIO CURRENTLY FOCUSES ON THE FOLLOWING FUTURE CAPABILITY VISIONS:

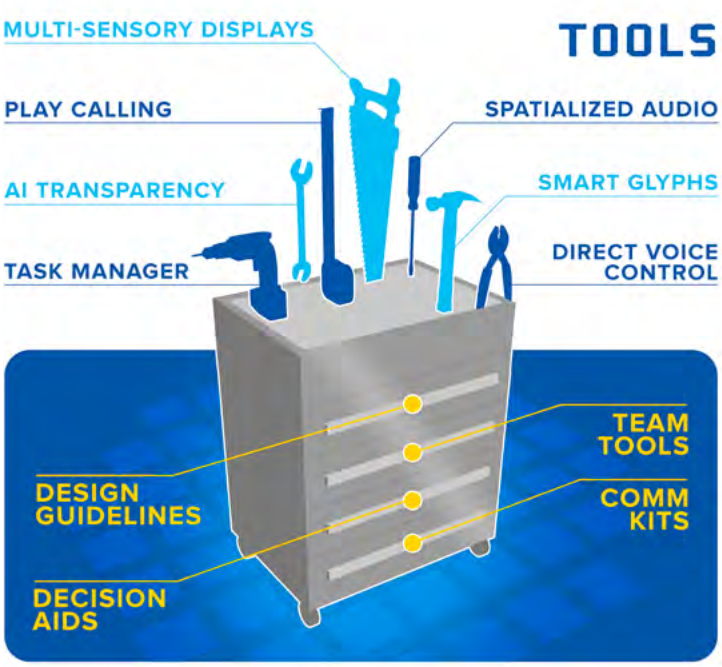
Joint All Domain C2

Manned-unmanned teaming

Data-driven ISR fusion

Emerging Space applications of warfighter interfaces

So, what does AWI contribute to the fight? We strive to provide a "toolset" of interaction and analytic solutions that empower Airmen and Guardians to accomplish their missions more rapidly and effectively, providing decision superiority to our forces in the next fight. Specific outputs/products of our research include operator interface design guidelines/knowledge, adaptive & intuitive multi-sensory interface concepts, communication enhancements, decision aiding & enriched sense-making methods, and teaming assessment and collaboration solutions.



Graphic by Mr. Will Graver

With the hope of a return to normal operations, the AWI CTC is poised to execute targeted research agendas along several fronts to address the highest-priority needs of the USAF and DoD. With new initiatives beginning in the areas of JADC2, Space applications, and Cyber operations, FY21 looks to be an exciting and productive year for the adaptive warfighter interface research and design! ☆

Dr. Mark Draper, Adaptive Warfighter Interfaces Core Technical Competency Lead and Principle Engineering Research Psychologist, 711 HPW/RHW

"THE AWI CTC IS POISED TO EXECUTE TARGETED RESEARCH AGENDAS ALONG SEVERAL FRONTS TO ADDRESS THE HIGHEST-PRIORITY NEEDS OF THE USAF AND DOD."

— Dr. Mark Draper
Adaptive Warfighter Interfaces Core Technical Competency Lead
and Principle Engineering Research Psychologist, 711 HPW/RHW

ACCOMPLISHMENTS:

- **Developed, deployed, and tested the novel Dialogue Experiment Testbed (DiET) for communications research and corpus collection.**
- **Transitioned dismounted communication system specifications for capabilities including multichannel recording, spatial audio, radio replay, and auditory-speech display into requirements for the upcoming Special Warfare Assault Kit (SWAK) acquisition plan.**
- **Two SMART crew scheduling and forecasting interface technology efforts were approved as part of official Air Mobility Command system requirements.**
- **Extensive interviews and cognitive task analyses have been conducted across several warfighter communities, including C2 operations, USAF special operations ATAK users, multiple intelligence analyst & intelligence operations organizations, tanker operations, cyber operations, security forces, etc.**
- **Prominent "human effectiveness" roles established to support two exciting, new AFRL 6.3 Programs, the Fight Tonight Agile Planning Program and the Rocket Cargo for Agile Global Logistics Vanguard.**
  - The AWI CTC will identify key human touchpoints and cognitive requirements while also developing and evaluating design solutions for meaningful human engagement with these revolutionary systems.
- **Memorandum of Understanding (MOU) signed between 711 HPW/RH and the Defense Language Institute for a 'light' version of our Haystack Multilingual Multimedia Information Extraction and Retrieval System.**
- **Initiation of exciting new Seedling research that will:**
  1. Develop real-time, objective measures of trust in distributed teams by assessing how trusting a teammate influences attention, cognitive efficiency, and performance.
  2. Explore the detection and mitigation of "auditory deepfakes" embedded in communication channels.

SYSTEMS ANALYTICS

Dr. Vincent Schmidt
Systems Analytics Core Research Area Lead,
711 HPW/RHWA

Even revolutionary technologies may inhibit, rather than enhance, operator performance when poorly implemented. Recognizing this critical component of success, the Systems Analytics CRA remains focused on accelerating the design and assessment of mission-relevant, *Airman-centric* data analytics capabilities at speed and scale. Over the past year, the research teams prioritized building upon core strengths:

- Enduring strategic alignment with DoD, AF, and AFRL priorities, including support for AI/ML analytics, Joint All Domain Operations, Space, Next Generation ISR, Advanced Information Warfare, etc.
- A foundational Cognitive Systems Engineering approach to technology development uniquely positions CRA to lead the research and design of human-centered data analytics for high-uncertainty, high-consequence systems.
- Interdisciplinary research teams leverage diverse workforce expertise to combine perspectives and principles from Cognitive Engineering/Human Factors and technologies and techniques from Data Analytics/ Computer Science with in-house software development and Warfighter domain expertise to rapidly deliver agile technology solutions.
- Strong and sustained connections to operational customers provides direct knowledge of mission requirements and enables accelerated technology transition, as do ongoing partnerships with DoD and Academic research organizations.

This focus has created new opportunities for the CRA:

- Alignment with Rocket Cargo Vanguard demonstrates strategic value and increasing scope of impact.
- Growth in partnerships with external organizations driving DoD data analytics investments, including NGA and Defense Advanced Research Projects Agency (DARPA).
- Competition for new DoD resource streams targeted at maturing prototype capabilities to support mission operations, including establishment of new funding vehicles.

In sum, the Systems Analytics CRA remains well positioned to drive the development of mission-ready data analytics now and into the future. ☆

Dr. Daniel Zelik, Systems Analytics Core Research Area Lead, 711 HPW/RHWA

ANALYTIC-ENABLED COGNITION

Current AI/ML-driven analytics solutions manipulate data in ways that are both unobservable and unexplainable to users, rendering them a poor fit for operational environments that are highly-cognitive, tightly-coupled, and rely on rich and specialized expertise to achieve mission success. The Analytic-enabled Cognition (AeC) line of effort seeks to quantify the impact of analytics on thinking and reasoning in order to tailor capabilities to the context-specific cognitive requirements of Airmen and Guardians. Over the last year, there was significant progress across all AeC lines of effort. Multiple analytics studies were conducted, leading to the submission of several journal articles. Meaning-Making in the Information Environment technologies supported operational exercises, and a new Cognitive and Physiological Performance assessment effort launched. In addition, one legacy set of projects concluded, enabling the start of two new efforts: Conversational AI and Insider Threat studies. ☆

Dr. Daniel Zelik, Systems Analytics Core Research Area Lead, 711 HPW/RHWA

> **The Analytic-enabled Cognition line of effort seeks to quantify the impact of analytics on thinking and reasoning in order to tailor capabilities to the context-specific cognitive requirements of Airmen and Guardians."**
>
> **— Dr. Daniel Zelik**
> Systems Analytics Core Research Area Lead, 711 HPW/RHWA

TASK 1: SYSTEM ANALYTICS ASSESSMENT (SAA)

SUB-TASK 1.1: ANALYTICS INTEGRATION FOR COGNITIVE PERFORMANCE

The Cognitive and Physiologic Performance (CPP) program has started its first of three years. The program will validate Basic Performance Resources (BPRs) for analyst candidate selection and develop the real-time assessment tools for identifying the current cognitive activities of the ISR airman in order to optimize the AI assistance. CPP was originally intended to focus on the activities and future aids for individual 1N1 imagery analysts. However, the Air Force Distributed Common Ground System (AF DCGS) is currently undergoing a shift from its traditional platform-centric exploitation model with individual analysts assigned to individual platforms with operations based on Analysis and Exploitation Teams (AETs). Each team is expected to use Multi-INT data sources to think holistically to creatively address problems.

Consequently, to maximize the payoff for this emerging model of operations, CPP is focusing on future DCGS operations and has commenced a Cognitive Task Analysis (CTA) to assess AET operations and requirements. The results of the CTA will be leveraged to identify the cognitive processes most relevant to AET analysts in general. Based upon these analyses, real-time neuro-ergonomic indices will be selected and tested for sensitivity to the targeted cognitive states. The studies will be initially conducted in a variety of task environments to fully test the robustness of the neuro-ergonomic indices, but the research will ultimately transition to validation tests in the Analyst Test Bed (ATB) facility that's reconfigured to simulate AET operations. ☆

Dr. Michael Vidulich, Research Psychologist, 711 HPW/RHWAR

Screengrab by 711 HPW/RHW

SUB-TASK 1.2: TACTICAL ANALYTICS STUDIES

For intelligence analysts, meaning-making (sensemaking) is challenging because it frequently involves making inferences about uncertain data. One way to enhance meaning-making may involve collaboration from a machine agent, such as one that uses AI algorithms to direct analysts' attention to people and vehicles in scenes. The System Analytics Assessment Team determined whether the human capability for meaning-making, or for identifying essential elements of information (EEIs), could be enhanced either by a simulated Project Maven or by a simulated agent that directed attention to EEIs. Full motion videos of simulated compounds viewed by an overhead camera were created. Meaning-making was assessed by participants' ability to predict events and identify signs.

Novices and intelligence analysts participated, and the results showed that EEI identification was enhanced by having the simulated agent direct the participants' attention to the EEIs, but not by directing attention to people and vehicles, as implemented by many current approaches. Participants' meaning-making was not enhanced by either type of machine agent; subsequently, guiding attention to features in a scene enhances their identification, whereas indiscriminate steering of attention to entities in the scene does not improve understanding of the holistic meaning of events. These results contribute to our goal of determining which human-machine systems improve the meaning-making capability of intelligence analysts in the field. ☆

Dr. Robert Patterson, Senior Psychologist, 711 HPW/RHWAR

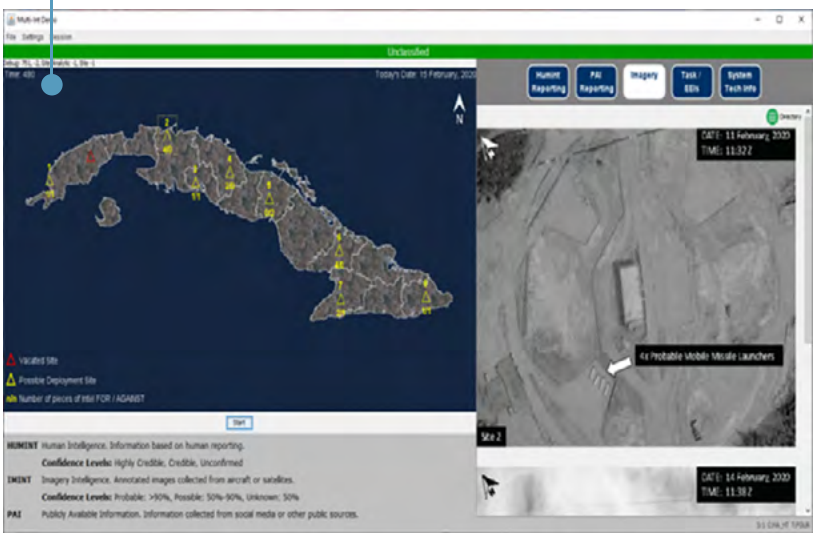
SUB-TASK 1.3: OPERATIONAL ANALYTICS STUDIES

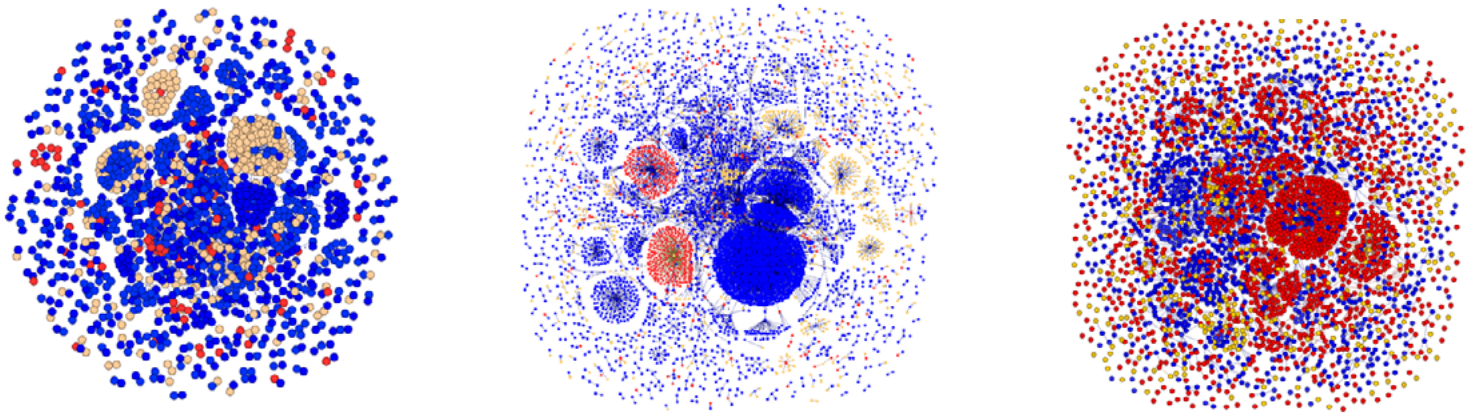
The Multi-INT Analytics study investigates how decision analytic transparency and accuracy influence the way in which people integrate multiple types of intelligence information to make a cohesive assessment of the situation. The scenario for a study currently underway, developed with input from National Air and Space Intelligence Center (NASIC) analysts, tasks people with determining the most likely site of a relocated Integrated Air Defense System. People are given access to personal intelligence data (PAI) data (e.g., Twitter® feeds, Facebook® posts), overhead imagery, and human intelligence (HUMINT) reporting. Data collection for the first study is nearing completion, with almost half of the data collected remotely via WebEx. Results will be analyzed in the upcoming months and will feed into the design future studies.

Conversational Artificial Intelligence (CAI) includes analytics that translate user verbal or textual input into actionable machine analytics, giving a user some type of output. Siri, Alexa, and Cortana are some examples of CAI that have been developed for the general user population to solve simple tasks. The Information Directorate is developing a CAI as part of their Multi-Source Analytics for Conversational Intelligence (MACI) program for use by intelligence analysts. MACI will integrate with much more complex analytics to provide input for analysts when solving Requests for Information (RFI). Our group is researching how the introduction of CAI systems impacts human cognition when it is used for complex tasks in high-stakes environments, with results feeding into design considerations for the MACI program. ☆

Ms. Kellie Turner, Research Psychologist, 711 HPW/RHWAR

Screengrab by 711 HPW/RHW





Graphic by Dr. Kathleen M. Carley

TASK 2: MEANING-MAKING IN THE INFORMATION ENVIRONMENT (MMIE)

SUB-TASK 2.1: INFLUENCE CHARACTERIZATION

Operating in the information environment today is highly challenging for military warfighters. Due to the difficulty in assessing influence in the Information Environment (including traditional and social media), a Social Digital Media playbook was developed to help characterize actions, events, and communications (Beskow and Carley, 2019) based on a new Information Maneuver (BEND) framework that characterizes influence in terms of information maneuvers, both targeting communities and content or control of the narrative. Much of the focus in developing analytics has focused on text content; however, visual (videos, memes, etc.) and audio (including audio chat) are emerging as important mechanisms for communications, influence. New analytics can assist planners, information operators, intelligence analysts with adaptive planning, triggering, and cueing of sensors, strategic communication, etc. Further, by understanding the online behaviors and mechanisms of influence, it can help forecast behaviors offline (e.g., civil unrest leading to protests) which could lead to better military intervention strategies. Another gap exists in the ways operational individuals are trained to understand mechanisms of influence in the Information Environment. This has huge ramifications for Force Security, Operations Security, Intelligence Surveillance and Reconnaissance operations, and Information Operations.

A prototype system for video-based OIE characterization was developed and made available online. The system includes innovative techniques, such as barcode-based characterization of videos, a technique that garnered a conference best paper (Erol, R., Rejeleene, R., Young, R., Marcoux, T., Hussain, M.N., and Agarwal, N., "YouTube Video Characterization Using Moviebarcode," Proceedings HUSO 2020 : The Sixth International Conference on Human and Social Analytics, 15-19). The system also includes topic modeling-based characterization, Plutchik's opinion-based characterization, toxicity/hate speech-based characterization, and content diversity-based (information theoretic measures) based characterization of visual space. Ongoing work is focusing on experimenting with emotions and attitude extraction from podcasts and audio from YouTube based on speech intonations. This approach will be fused with text-based and color theory-based emotion extraction to enable a more accurate assessment.

"Information in its many forms has become a significant component of national power—the primary medium of competition between the United States and its adversaries."

— Mr. Andrew Milburn and Ms. Shawna Sinnott
mwi.usma.edu

Another research effort under this task is focused on developing an BEND game named "Operational Mastery of the information Environment (OMEN) training individuals, including synthetic data that represents the ever-evolving behaviors/tactics in the Information Environment, canonical actors, including different types of bots, cyborgs and their roles. This includes developing canonical bot types, troll types, and cyborg (switching between human/bot) types, developing (BEND sequence) maneuver types with typical activity for various actor types, and identify sets of counter maneuvers. For example, bots can be megaphones (amplifying messages/narratives/content) or brokers (bridging networks), functionally supporting both **build** and **bridge** maneuvers in order to amplify conversations and communities. Bots can promote conspiracies, help to spread disinformation, distort a conversation by hashtag latching, or distort an identity by brandjacking. Bots (deep fakes) or trolls can also distort or manipulate opinions. New analytics for detection of hate speech are important, along with network metrics that point to a network becoming more like an "echo chamber," for detecting increasing levels of radicalization that can spill over into offline protests, violence (e.g., the heated/polarized "re-open" conversation in many places). New BotHunter, TrollHunter, and Hate Speech detection analytics have been added to the existing ORA and NetMapper toolkits developed by Carnegie Mellon University/Netanomics. ☆

Ms. Laurie Fenstermacher, Principal Electronics Engineer, HPW/RHWA

SUB-TASK 2.2: RESILIENCE TO MISINFORMATION

One new MMIE in-house research effort focuses on auditory deepfakes. An auditory deepfake is when a real person's voice is "cloned" to produce synthetic audio data, to "Photoshop" another's voice. This has implications for national security with regards to OPSEC. For example, an adversary could use this technology to impersonate a commander. As platforms become less text-based, auditory deepfakes could be the next tool for disinformation wars on social media. One example where this tool might be used is the emerging platform Clubhouse, which is an exclusively audio-based app where people interact in live conversations. To this end, Dr. Kathleen Larson and colleagues developed a research project that examines the intersection of synthetic voices and disinformation. The summer interns and cadets in RHWAR created experimental stimuli by recording their real voices and synthetic voices (from a deepfake generator), vocalizing real and fake headlines and excerpts. To examine the interaction of familiarity, the interns distributed a stimuli assessment on family members and close friends. While listeners were less likely to fall for the synthetic versions, content read by synthetic voices were also perceived as less truthful. These stimuli will be used and expanded upon in future experiments that will examine multiple factors that influence the human detection of deepfakes, such as familiarity, gender, affect, and context. The results of these studies will provide insights on the situations where the Warfighter might be the most vulnerable to synthetic voice tactics. ☆

Dr. Kathleen Larson, Research Psychologist, 711 HPW/RHWAR

Photo by Airman 1st Class Alexcia Givens
Graphic by Ms. Shania Horner

TASK 3: JOINT ALL DOMAIN INTEGRATED ISR RESEARCH (JADII)

The JADII team continues to amalgamate academic research with current and future operational joint air and space warfighting requirements. Two notable recent efforts include investigations of use of online survey tools to provide analytic insight (Sphinx) and trust in AI to support a planning task (path planning decision support study).

Under the first effort, the Sphinx tool from past reporting is undergoing major surgery. While the underlying methodologies remain consistently tied to structured analytics and crowd sourcing, Sphinx currently uses the analytic technique known as the method for defining analytical questions (MDAQ) to diagnose the problem statement. Depicted as a structured approach, the MDAQ guides analysts to assess, process, hypothesize, and deliver a decision with high confidence (Berger et al., 2009). Moreover, metacognition and critical thinking efforts led educators to the concept of scaffolding techniques (James, I. & Okpala, C., 2010), which help instructors and technical writers shape information that improves understanding and academic performance. Scaffolding techniques generally consist of problem representations, generation and selection of solutions, justifying the selection, and then testing and evaluating the result. The purpose of this effort is to determine whether the MDAQ or Scaffolding methodology is beneficial for facilitating problem solving, and compares these methods to a control condition when gathering EEIs from a mystery narrative.

A total of 175 participants received instructions to read and solve a mystery story adapted from an online mystery solving website, 5minutemystery.com. Reading comprehension level assessments conducted on each mystery also contained embedded EEIs. Random assignments given to each participant consisted of either the control, MDAQ structure, or scaffolding technique. In addition, stimuli validation testing occurred on multiple mystery stories.

The findings provided underlying evidence that implementing a scaffolding-structured analytic approach did significantly improve performance ($p < 0.01$) compared to the MDAQ and the control analytic approaches. The results from the study provided evidence that structured analytic approaches can improve performance when information is vague and requires additional interpretation compared to a control condition. Participants directed toward answering the narrative with a scaffolding analytic approach improved the collection, interpretation, and dissemination of potentially overlooked, critical EEIs. Further research is required to determine how to best optimize these structured analytic approaches, and how they can transition into a military ISR environment to support

Continued on Next Page...

the Warfighter. Expanding this research to test stimuli more closely for alignment to the Sphinx model continues.

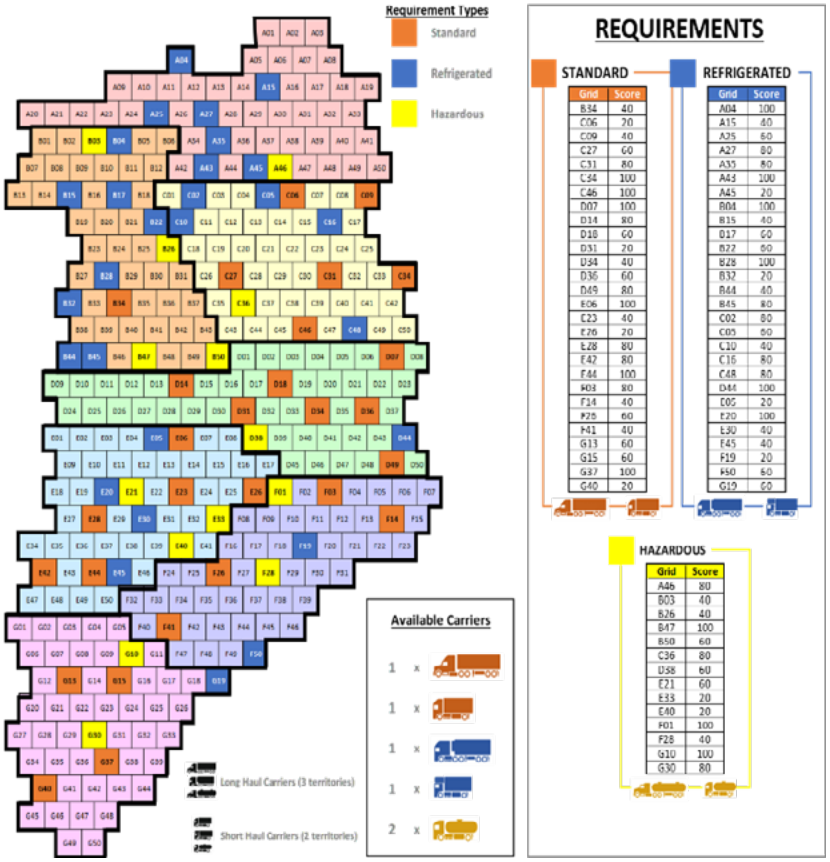
Pivoting to the second, the military increasingly relies on Decision Support Systems (DSS) and other tools to recommend courses of action and augment Warfighter performance on critical tasks. This is particularly important for path planning operations, where planners must consider complex tradeoffs and contingencies based on available assets, distance, environment, resources, threats, and target priority. To replicate this task in a more general-applied context, JADII researchers developed a path planning task that emulated long-range trucking dispatch management. Participants provided a quality control check of four simulated DSS recommendations, which ranged from perfect (100%) accuracy to subpar (40%) accuracy. Individuals were also presented with different types of explanations for how the algorithm would determine where trucks should be allocated. Participants reported lower trust of lower accuracy DSS. However, their quality control performance was significantly lower when the DSS was below perfect accuracy. This demonstrated that while participants were able to successfully calibrate perceived trust in their DSS, they nevertheless experienced performance decrements, possibly due to anchoring on the DSS's incorrect result. Furthermore, participants reported the greatest understanding, strongest trust, and highest preference for the simple written explanation of the DSS's algorithm over more complex presentations and descriptions.

The findings of this study provide the groundwork to understand the relationship between automation-reliance, trust, and performance, to determine when it is appropriate to allow automation to make and contextualize recommendations to Warfighters in operational environments. ☆

Dr. Jerred Holt, Research Scientist, 711 HPW/RHCMD

> **...The military increasingly relies on Decision Support Systems (DSS) and other tools to recommend courses of action and augment Warfighter performance on critical tasks."**
>
> **— Dr. Jerred Holt**
> Research Scientist, 711 HPW/RHCMD

Screengrab by 711 HPW/RHW



SOLUTION									
Truck Type	**Trailer Type**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**Assigned Regions**
1 x Long Haul	Standard			X	X	X			C&D&E
1 x Short Haul	Standard						X	X	F&G
1 x Long Haul	Refrigerated	X	X	X					A&B&C
1 x Short Haul	Refrigerated				X	X			D&E
1 x Short Haul	Hazardous	X	X	X					B&A or C
1 x Short Haul	Hazardous						X	X	F&G

Required Trailer		**A**	**B**	**C**	**D**	**E**	**F**	**G**
Standard	Count	0	1	6	7	6	4	4
	Score	0	40	400	440	420	220	240
Refrigerated	Count	7	8	5	1	4	2	1
	Score	480	500	340	100	200	80	60
Hazardous	Count	1	4	1	1	3	2	2
	Score	80	240	80	60	100	140	180

SENSE-MAKING AT SCALE

Operators across missions must rapidly make meaning from "Big Data" and overcome the inherent complexity and uncertainty of multi-dimensional data to create actionable insights in real time. The Sense-making at Scale (SaS) line of effort seeks to design analytics to enhance sense-making and mitigate data overload in order to enable Airmen and Guardians to rapidly extract meaning from complex, uncertain, multi-dimensional data sources. SaS LOE projects have progressed via technology demonstrations, prototypes, and customer engagements. In addition, eXplainable Artificial Intelligence (XAI) and Applied Operational Analytics (AOA) effort have merged into a single Applied eXplainable Analytics (AXA) task to better align mission focus across projects. ☆

Dr. Daniel Zelik, Systems Analytics Core Research Area Lead, 711 HPW/RHWA

> **The Sense-making at Scale line of effort seeks to design analytics to enhance sense-making and mitigate data overload in order to enable Airmen and Guardians to rapidly extract meaning from complex, uncertain, multi-dimensional data sources"**
>
> **— Dr. Daniel Zelik**
> Systems Analytics Core Research Area Lead, 711 HPW RHWA

TASK 4: DYNAMIC WIDE AREA DISCOVERY AND EXPLORATION (DWADE)

The Dynamic Wide Area Discovery and Exploration (DWADE) line of effort is built around work domains requiring practitioners to create insights from immense datasets. DWADE is comprised of three projects, each of which seek to support practitioners confronted with data overload. The cognitive work performed by successful practitioners in each of these domains share many similarities. As a result, the larger goal is to develop generalized principles for building tools to support practitioners confronting data overload.

DWADE's first project, Hound, seeks to build support tools for intelligence analysts confronting data overloaded from ever-increasing sensor coverage and sensor diversity. Legacy tradecraft for Geospatial Intelligence exploitation does not scale with the growing availability of data. Hound helps traditionally siloed analysts discover new opportunities for data exploitation by correlating multiple types of intelligence data into tipping and cueing. Using this bottom-up approach to tradecraft, also known as Activity Based Intelligence (ABI), will allow analysts to be more efficient in how they spend their limited manpower and help discover activity that no one knew to look for.

DWADE's second project, Flatline, supports communication intelligence (COMINT) analysts who monitor and translate radio communications over large areas. Flatline provides sensor metadata visualization to help analysts make more effective decisions about which frequencies to monitor. The effectiveness of Flatline's data visualizations is currently being investigated by Ohio State University's Cognitive Systems

Continued on Next Page...

DWADE'S 3 PROJECTS

Hound

Seeks to build support tools for intelligence analysts confronting data overloaded from ever-increasing sensor coverage and sensor diversity.

Flatline

Supports COMINT analysts who monitor and translate radio communications over large areas.

Vector

Ingests the output of AI/ML algorithms designed to track objects in video.

Engineering Lab (CSEL). The outcome of this study will help researchers develop generalized principles for supporting human decision making at scale.

DWADE's third project, Vector, originally started as a tool being developed by the National Geospatial Intelligence Agency (NGA). Vector is a human-facing workflow tool that ingests the output of AI/ML algorithms designed to track objects in video. Metadata about that output is visualized to intelligence analysts. Analysts can use those visualizations to quickly and efficiently stitch that algorithm output into full, decision-quality motion tracks. The current generation of tracking algorithms is highly prone to broken tracks or other forms of error in situations where human analysts are very resilient. The effectiveness of Vector's kinematic metadata visualization in supporting analyst decision making is currently being evaluated in a joint effort between RH and NGA Research. The methodology of this study will become a generalizable test-harness for conducting similar evaluations on different technology in the future.

These three projects all require practitioners to develop insights while being confronted by data overload. By understanding the common cognitive work required by all three, DWADE will develop generalizable design principles that can be applied across Air and Space Force missions. ☆

Dr. Taylor Murphy, Cognitive Systems Engineer, 711 HPW/RHWAR

TASK 5: APPLIED EXPLAINABLE ANALYTICS (AXA)

SUB-TASK 5.1: OPERATIONAL ANALYTICS

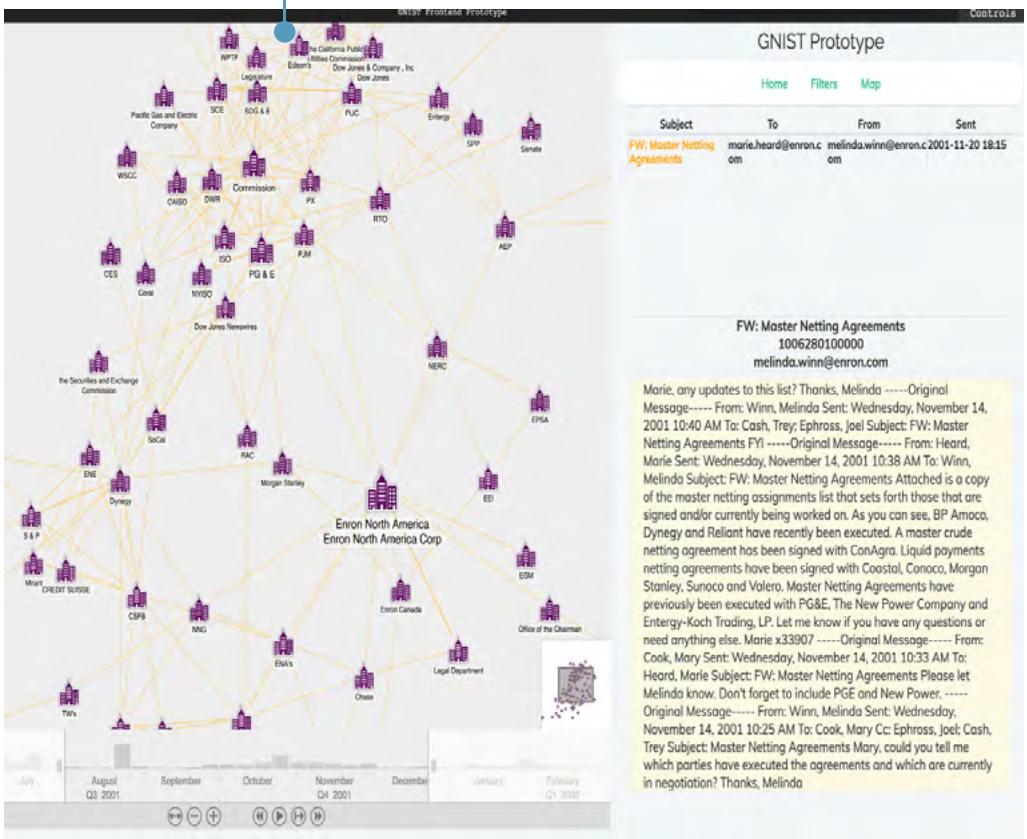
Generalized Network Insights from Situated Text (GNIST) looks at how a combination of analytics and user-centered visualizations can best enable rapid and calibrated characterization and sense making of complex networks. GNIST operates on unstructured text to discover relationships and is capable of handling 50+ languages to perform NLP, including topic modeling and name entity recognition. It displays 100+ nodes, their interactions, and temporal updates in near real-time while reducing network dimensionality by combining subnetworks and structured filters. Research is being conducted to investigate whether the type of network analytic (Topic Modeling vs Knowledge Graph) facilitates different levels of understanding (e.g., literacy) of a network and if the type of task (e.g., investigate a specific person/activity vs. a general descriptor) interacts with network representation type. ☆

Ms. Kellie Turner, Research Psychologist, 711 HPW/RHWAR

By understanding the common cognitive work required by all three (projects), DWADE will develop generalizable design principles that can be applied across Air and Space Force missions."

— Ms. Taylor Murphy
Research Engineer, 711 HPW/RHWAR

Screengrab by 711 HPW/RHW



SUB-TASK 5.2: APPLIED TOPOLOGICAL DATA ANALYSIS (TDA)

The Applied Topological Data Analysis sub-task has two key focus areas: development of an "XAI Silo" and development of a government-owned framework for topological analysis (ATOM). XAI Silo is an application to host and enable rapid development, displays, and visualizations of eXplainable Artificial Intelligence. This application, currently under development, already has the capability to host several implementations of novel XAI research and several examples of traditional machine learning techniques with a healthcare focus. XAI Silo will give collaborators the opportunity to quickly use the functionality developed in the eXplainable Artificial Intelligence Laboratory. Technology developed for customer-specific use cases, such as Suicide Risk Prevention and Air and Space Force Education Training Command, will have an easy-to-install application to use XAI to make informed decisions.

The goal of ATOM is to implement an algorithm called the "Topological Hierarchical Decomposition (THD)." The THD works by recursively reducing data into two dimensions and then clustering on the lower-dimensional representation. The previous implementation the THD required commercially licensed software from a company called Ayasdi and also an internet connection. The ATOM developer has successfully implemented a zero-cost and offline THD. The THD is an eXplainable Artificial Intelligence algorithm that has been demonstrated as a potent stand-alone algorithm. It has also been used with great success as a feature engineering method aiding traditional machine learning techniques. ☆

Mr. Ian Joyce, Research Data Scientist, 711 HPW/RHWAI

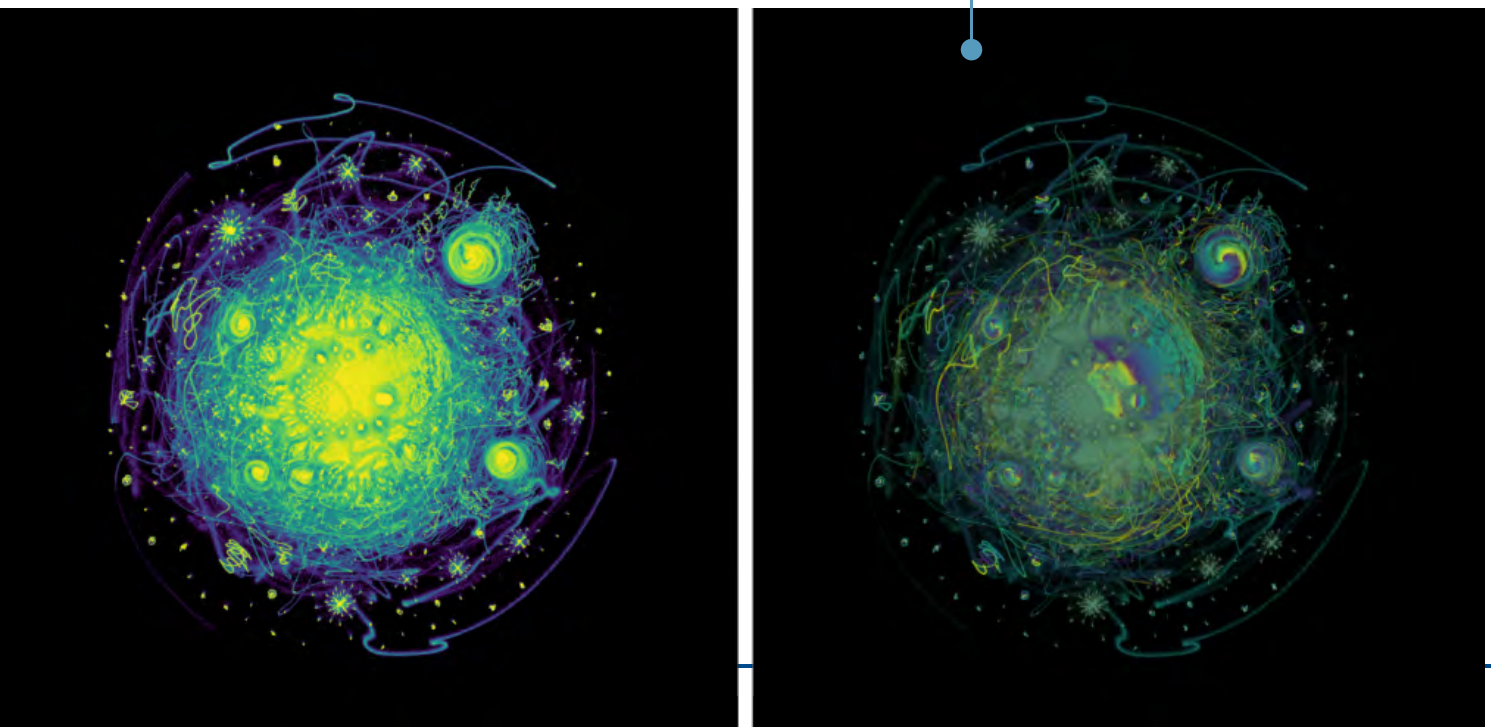
TWO KEY FOCUS AREAS:

Development of an "XAI Silo"

Development of a government-owned framework for topological analysis (ATOM)

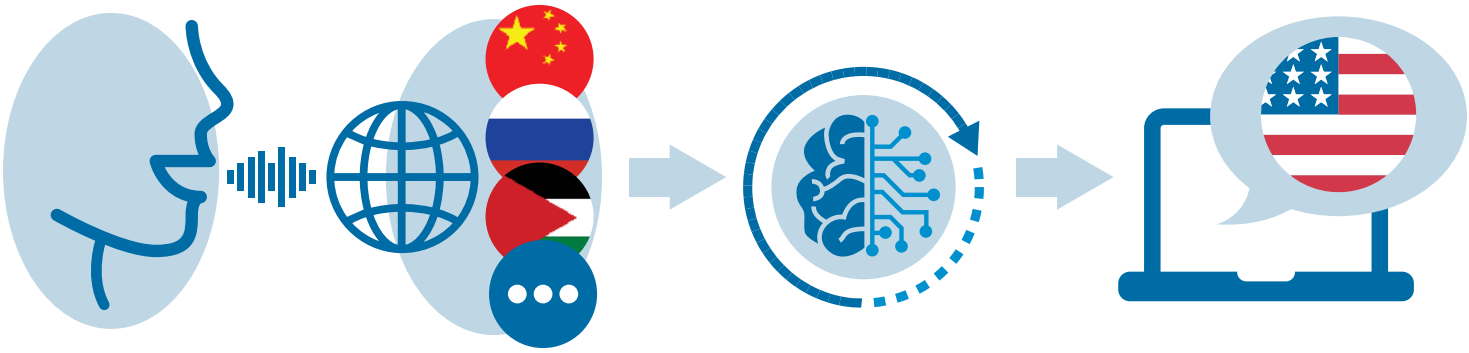
> **[Topological Hierarchical Decomposition] has also been used with great success as a feature engineering method aiding traditional machine learning techniques."**
>
> **— Mr. Ian Joyce**
> Research Data Scientist, 711 HPW/RHWAI

Screengrab by 711 HPW/RHW



A complex dataset reduced to two dimensions to reveal its underlying structure.

Graphic by Ms. Shania Horner



AUTOMATIC SPEECH RECOGNITION

NATURAL LANGUAGE PROCESSING

MACHINE TRANSLATION

TASK 6: HUMAN LANGUAGE TECHNOLOGY (HLT)

Human Language Technology (HLT) works on technologies for processing human language found in various media files. While there are a variety of technologies that the group works on, the major focus areas have been automatic speech recognition (ASR) and machine translation (MT). ASR automatically provides a text transcript for an audio track, while MT automatically translates text from one language to another. The group conducts research and develops larger systems that incorporate various HLTs.

Over the past year, significant advancements in ASR and MT system performance have occurred. In ASR, a state-of-the-art deep learning neural-based architecture has been implemented that has shown significant reductions in error rates across three different languages. This new model architecture is currently being expanded into additional languages, and there is a plan to explore combination outputs of this Convolution-Augmented Transformer network with our existing factorized time-delay neural network to see if this will further improve ASR performance. Advancements in MT have revolved around augmenting the training data by normalizing the source language training text to look like ASR output text (no punctuation, lowercased text, numbers and symbols written as spoken) and also by adapting these general-purpose MT systems to specialized domains of military interest, such as cyber or technical journals, without requiring additional training data. By augmenting the training data and tuning models to specific domains, sizeable performance improvements have been achieved. In addition, transformer-based punctuation models have been developed to re-punctuate ASR output and automatically identify sentence boundaries when needing to translate a video or audio file, and this has produced significant improvements to the readability of the MT output.

These recent advancements in ASR and MT have been used by the HLT Group to participate in international evaluations such as the International Conference on Spoken Language Translation (IWSLT) [iwslt.org] and the Workshop on Statistical Machine Translation (WMT) [aclanthology.org/venues/wmt]. Papers describing, in detail, the work highlighted above have been published to each of these Workshops. In addition to publications, all of these advancements will be or have been incorporated into the Haystack system–which is under the Analytic Tools Product Line and has already been deployed to multiple customers. ☆

Dr. Raymond Slyh, Principal Research Electronics Engineer, 711 HPW/RHWAI

WHAT IS AUTOMATIC SPEECH RECOGNITION (ASR)?

ASR automatically provides a text transcript for an audio track.

WHAT IS MACHINE TRANSLATION (MT)?

MT automatically translates text from one language to another.

COLLABORATIVE INTERFACES AND TEAMING

Dr. Chris Brill
Collaborative Interfaces and Teaming Core Research Area Lead, 711 HPW/RHWC

The Collaborative Interfaces and Teaming Core Research Area (CRA) focuses on the research and development of flexible, directable, and transparent human-autonomy teaming solutions; the science of human teams in distributed multi-domain contexts; assessments of fielded autonomy systems; and the development of technologies to facilitate shared authority of autonomy and common ground within and between mixed human-autonomy teams. The Collaborative Interfaces and Teaming CRA's research priorities generally fall under two major themes: Human-Autonomy Collaboration and Distributed, Heterogeneous Teaming Solutions. ☆

Dr. Chris Brill, Core Research Area Lead, 711 HPW/RHWC

Understanding the design principles for adaptive collaborative interfaces for human and mixed human-autonomy teams is absolutely central to manifesting future Air Force capabilities that will ensure our edge over peer and near-peer adversaries."

— Dr. Chris Brill
Core Research Area Lead,
711 HPW/RHWC

An example of an operator pointing, clicking, and dragging the cursor to change the swarm heading direction following a degradation (which was denoted by red assets in this experiment). The right-hand side of the image (enhanced here) shows the full 3-minute trial notification panel, where participants are given updates regarding degradation occurrence, targets collected, and when an input may be offered. Adapted from Hamdan, I., Capiola, A., Alarcon, G. M., Lyons, J. B., Nishimura, K., Sycara, K., & Lewis, M. (2021). Exploring the effects of swarm degradations on trustworthiness perceptions, reliance intentions, and reliance behaviors. Accepted for presentation at the Human Factors and Ergonomics Society 65th International Annual Meeting. Baltimore, MD.

HUMAN-AUTONOMY COLLABORATION

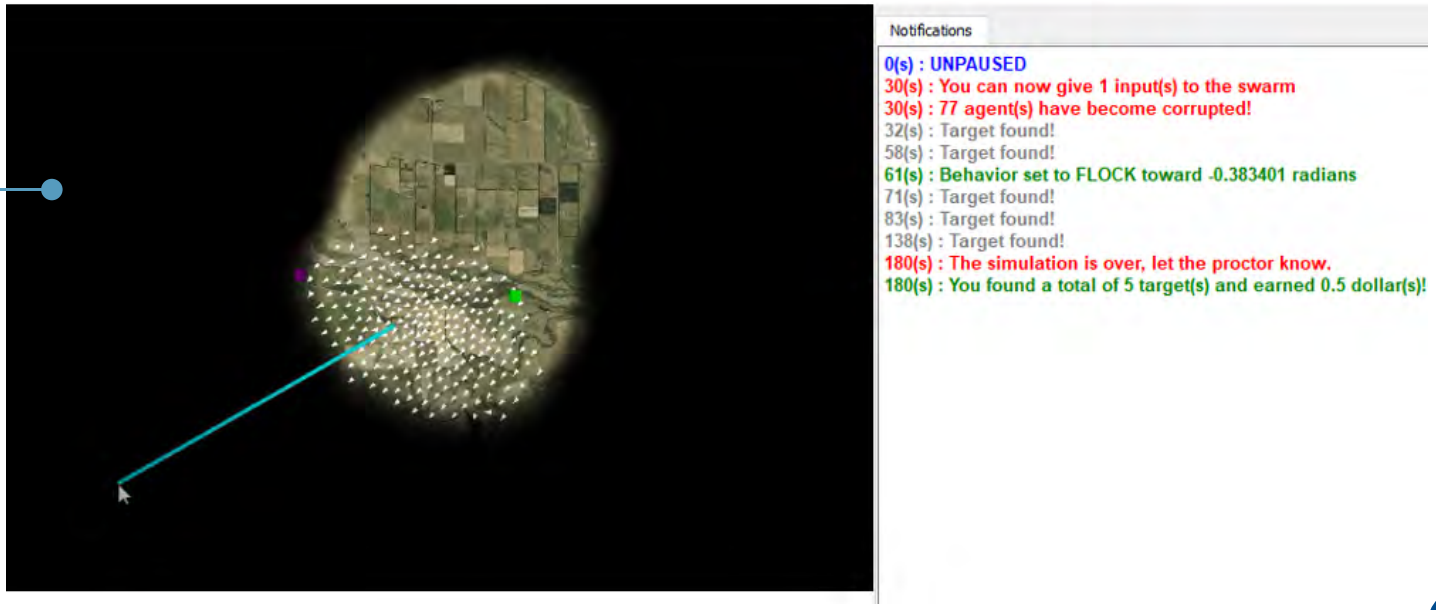
TRUST IN SWARMS

Swarms are comprised robots operating autonomously through local control laws. Swarm performance is difficult to assess, making appropriate levels of trust toward swarms critically important especially when the systems behave unexpectedly or are subject to countermeasures. As part of the Center of Excellence for Trusted Human-machine Teaming—comprised of AFRL researchers and academic stakeholders from Carnegie Mellon University—we have leveraged and augmented a testbed to investigate the effects of robotic asset degradation on the perceptions and intentions humans have in regards to relying on swarms. Our preliminary research from FY 20 showed that human trust toward swarms varied based on the proportion of asset degradations, suggesting that degradations to one facet of a system pervade into human beliefs about the entire system that support the notion of system-wide trust in human-swarm interactions. In FY 21, we addressed several limitations from our initial data collection. Specifically, our work was accepted for publication that assessed the relationship between human perceptions and behavioral trust (i.e., takeover redirection) of swarms. We also collected data from a large online sample to determine participants' ability to detect degraded assets in foraging scenarios, and we followed-up with another data collection we are currently analyzing. We are investigating the granular visual accordances participants' may use that potentially qualify the influence of asset degradation on degradation detection. Soon, we aim to launch our next iteration, investigating the effects of asset degradation on trust toward swarms controlling for swarm performance, automation cues, and granular features instantiated in our past work. ☆

Dr. August Capiola, Research Psychologist, 711 HPW/RHWCT

> **Soon, we aim to launch our next iteration, investigating the effects of asset degradation on trust toward swarms controlling for swarm performance, automation cues, and granular features instantiated in our past work."**
>
> **— Dr. August Capiola**
> Research Psychologist, 711 HPW/RHWCT

Screengrab by Mr. Izz Aldin Hamdan



CYBER PROTECTION TEAM PLANNING RESEARCH

Cyber Protection Teams (CPTs) are tasked with ensuring the security and protection of cyber terrain for operations to take place. In order to provide mission assurance, CPTs must correlate and integrate information regarding the systems and networks that are pertinent to the mission. This is a challenging and complex task that can have detrimental implications on the mission if not executed effectively. In order to support this challenge, our research focused on the design and evaluation of a work aid for CPTs to efficiently conduct relevant planning activities.

The work aid was designed by conducting knowledge acquisition activities with CPT planners across the DoD and then transforming that information into work aid design concepts. The work aid was evaluated against current CPT planning methods by operators with various levels of planning experience. Participants were tasked with planning a fictional mission, utilizing both the traditional planning method and the work aid method. Data was collected from objective metrics, including time and accuracy of task performance, along with subjective perception of workload. Participants' understanding of mission-relevant information was also assessed through a series of situation awareness (SA) questions. When compared to the traditional planning method, the work aid method posed significant improvements to task time, SA question response time, and perceived workload. The results of this research will be presented at the 17th International Conference for Cyber Warfare and Security in March 2022 and published in the conference proceedings. Findings will support cyber protection team activities throughout both the Air and Space Forces along with joint-force activities. ☆

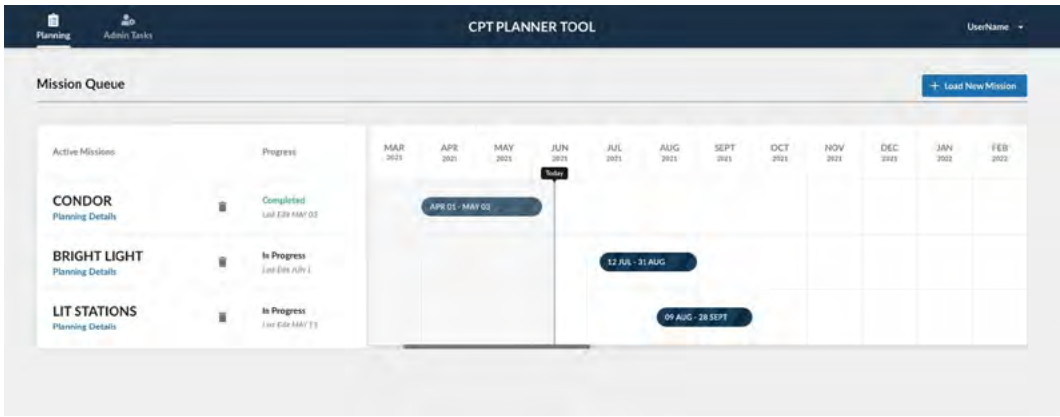
Dr. Kristen K. Liggett, Principal Human Factors Engineer, 711 HPW/RHWCI

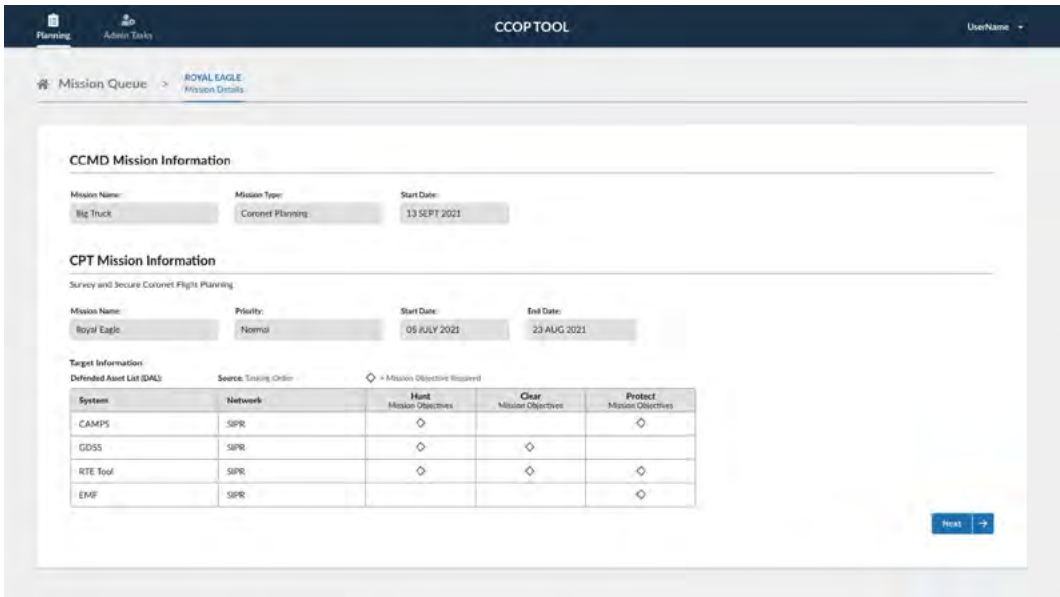
Dr. Gina Thomas, Research Psychologist, 711 HPW/RHWCI

1Lt Meghan Strang, Human Factors Engineer, 711 HPW/RHWC

HOW TO DESIGN A WORK AID:

1. **Conduct knowledge acquisition activities with CPT Planners across the DoD.**
2. **Transform information into work aid design concepts.**
3. **Develop a fictional mission for users to plan during the evaluation.**
4. **Evaluate usefulness and usability of work aid versus current CPT planning methods.**
5. **Incorporate changes from evaluation and user feedback into work aid design.**





Screengrabs by Ms. Ariel Swift

Photo by Airman 1st Class Ryan Lackey

MANNED-UNMANNED TEAMING:
INTERFACES AND METHODS

As machine-agent technology advances, the USAF will maintain its competitive edge in air defense by advancing the development and deployment of artificial intelligence (AI) and machine learning (ML) enabled agents across a variety of AF missions. For instance, agents may help pilots maintain situation awareness (SA) by displaying agent-teammate intentions and actions in a streamlined display. Such information can be most beneficial when a pilot is recovering from extreme events, such as gravity-induced loss of consciousness (G-LOC) or hypoxia. One of our goals for the AFRL is to evaluate how a Pilot's Associate (PA) may accelerate the recovery of pilot SA following a physiologically adverse event, such as G-LOC or hypoxia. The impact of these adverse events on pilot SA have not yet been investigated in the context of human-machine teaming where a supporting team of autonomous partners (APs) carry on a mission while the pilot is temporarily and unexpectedly incapacitated. Studying this will allow us to determine what information a pilot urgently needs from APs upon recovery from G-LOC; what impact this additional stream of data will have on pilot workload during recovery of SA; and how the availability of this additional data might interact with other contextual factors, such as air battle space complexity, along with moderate, rapid, and successful recovery of pilot SA. Findings from the experimental study will support the design of SA maintenance software and could be transitioned to the Skyborg Vanguard program. ☆

Dr. Greg Funke, Senior Research Psychologist, 711 HPW/RHWCT

Dr. April Rose Panganiban, Research Psychologist, 711 HPW/RHWCT

Dr. Michael Tolston, Research Psychologist, 711 HPW/RHWCT

TRANSPARENCY IN MACHINE LEARNING

Machine Learning (ML) is influencing many different aspects of modern society. However, ML is often considered a "blackbox," as the user does not know or understand the process the ML used to determine its assessment. Additionally, this "black box" aspect of ML leads to a lack of transparency for the user, which hinders trust. The current program explores "elegant failures" (i.e., the ML stating its confidence in its assessment) as an indicator of transparency and its ability to increase trust perceptions. ☆

Dr. Gene Alarcon, Research Psychologist, 711 HPW/RHWCT

WHAT IS G-LOC?

"A state of altered perception wherein (one's) awareness of reality is absent as a result of sudden, critical reduction of cerebral blood circulation caused by increased G force." –Pubmed.gov

WHAT IS HYPOXIA?

"Hypoxia is a condition in which tissues of the body do not receive sufficient oxygen (O2) supply." –Oxford Languages

TRUST EVALUATION OF AGCAS ON THE F-35

The Automatic Ground Collision Avoidance System (AGCAS) was fielded on the F-16 platform in 2014. Since then, it has experienced great success with more than 10 pilot saves to-date. Due to this success, AGCAS was fielded on the F-35 platform in 2019. However, the F-35 community immediately began to experience several unexpected activations of the AGCAS system following its fielding on the F-35. These unanticipated behaviors could create significant system distrust within the F-35 community, so the AFRL seeks to gauge pilot trust and understand the types of barriers to eventually trust in AGCAS within the community; subsequently, this study investigates the development of trust and reliance behaviors among F-35 pilots and will identify trust barriers for AGCAS in the USAF F-35 pilot community. Additionally, the study will identify interface issues, user experiences, concerns, policies, impact, and benefits of technology as they emerge. The results will support the F-35 Joint Program Office, Air Force Test Center, Office of the Undersecretary of Defense, and the Defense Safety Oversight Council (DSOC).☆

Dr. Chris Brill, Senior Research Psychologist, 711 HPW/RHWC

Ms. Anna Lee Van Abel, Associate Research Psychologist, 711 HPW/RHWCT

1Lt Jacob Ehrenstrom, Human Factors Engineer, 711 HPW/RHWC

Photo by Staff Sgt. Andrew Lee

COLLABORATIVE INTERFACES RESEARCH

For the future of warfighting, the importance of autonomy for realizing USAF employment of multiple manned and unmanned vehicle-teamed sensor platforms is well recognized. These new, mixed-initiative interactive systems must enable human-autonomy collaboration and teaming that pairs a human's pattern recognition and judgement capability with recent machine advances in artificial intelligence and autonomy in order to facilitate synchronized tactical operations. AFRL has made significant advances in developing intelligent agents and ground station command and control interfaces that will enable human-autonomy teaming to manage multiple assets. The approach utilizes a play-based delegation architecture and an intelligent Task Manager Interface to aid in workload management and shared team awareness. Instantiated in a high-fidelity system referred to as Intelligent Multi-UxV Planning with Adaptive Collaborative Control Technologies (IMPACT), this research capability (to-date) has been used to examine the utility of a wide spectrum of control for the management of 12 (simulated) and 6 (live) heterogeneous UVs by a single human operator teamed with an intelligent agent. This effort aims to evaluate various team structures and the interface and system designs that best support and improve teaming between multiple human members as well as between human-autonomy members. This involves developing new capabilities (such as intelligent agent technology and the simulation framework) and creating/enhancing display and control interfaces to support effective teaming interaction in the IMPACT station for multi-domain (including manned assets) command and control. Both part-task, lower-fidelity simulation/usability evaluations, as well as full-scale experimentation with the high-fidelity IMPACT system, are being used to evaluate alternative approaches to candidate features/designs of interfaces and decision aids. The results of the repeated cycles of design and evaluation will feed technical efforts that address future envisioned, complex mission scenarios. ☆

Dr. Elizabeth Frost, Research Psychologist, 711 HPW/RHWCT

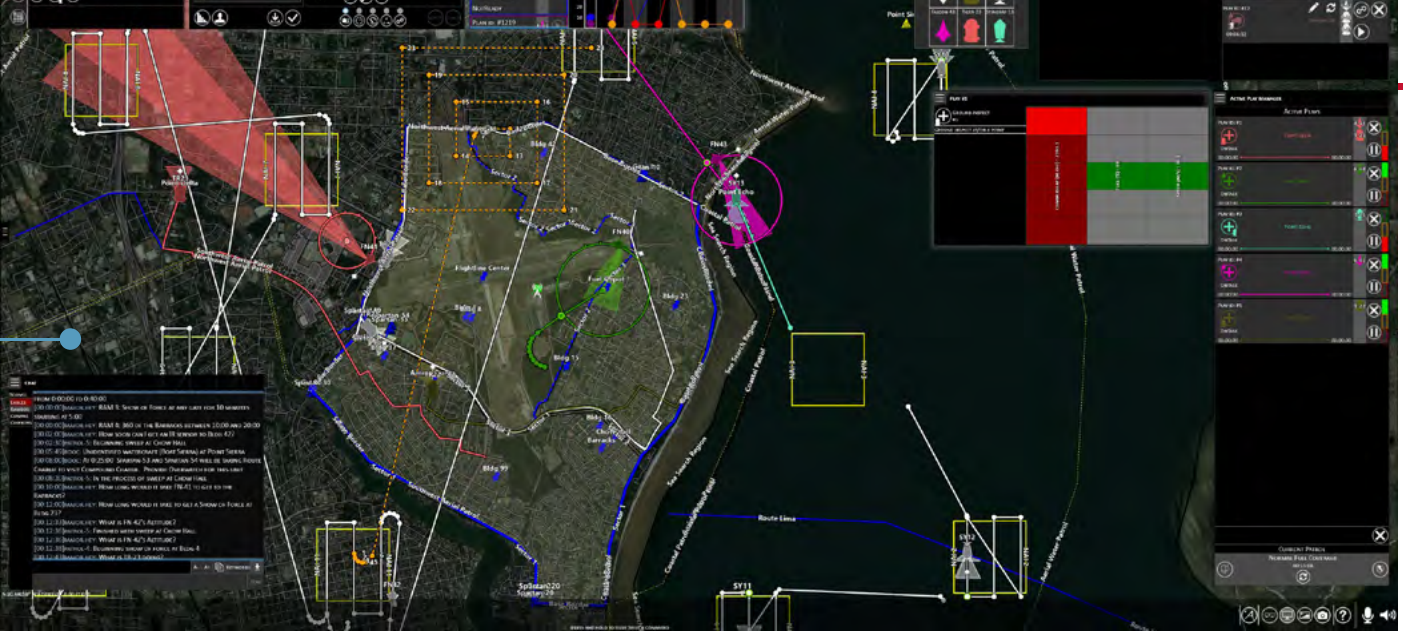
Ms. Jessica Bartik, Research Psychologist, 711 HPW/RHWCT

Dr. Hunter Rogers, Research Psychologist, 711 HPW/RHWCT

1Lt Jacob Ehrenstrom, Human Factors Engineer, 711 HPW/RHWC

Ms. Anna Lee Van Abel, Associate Research Psychologist, 711 HPW/RHWCT

Ms. Sarah Bowman, Computer Engineer, 711 HPW/RHWCI



Screengrab by Ms. Gloria Calhoun

THE SPREAD OF TRUST AND DISTRUST IN DISTRIBUTED HUMAN-AUTONOMY TEAMING CONSTELLATIONS

The work leverages previous research conducted at Clemson (Dr. McNeese), as well as research at Georgia Tech and the Air Force Research Laboratory (Dr. Gorman and Dr. Myers, respectively) on human-autonomy teaming (HAT), in addition to Freeman's work on Computer Supported Cooperative Work (CSCW). To prepare for the coming expansion and integration of human-autonomy teamwork, we planned a concentrated research effort that establishes a solid foundation to investigate the ability of trust to travel between teammates and teams in a human-autonomy teaming environment consisting of multiple HATs interacting – the spread of trust in multi-HAT constellations. The work consists of:

1. Repeated mixed-methods of empirical experiments that observe spread of trust within and between HATs.

2. The iterative development of a quantitative measurement specifically designed to capture and quantify the spread of trust between teammates within and between different teams.

3. An emphasis on ensuring HAT research is conducted in applied environments that are representative of those seen in DoD initiatives.

4. The landmark application of trust-repair techniques in a multi-HAT constellation with the goal of preventing organizational-wide spread of distrust.

This proposed research will contribute fundamental knowledge to the general scientific community studying human-autonomy interactions, applied human-autonomy teaming domains, and current and future initiatives of the Department of Defense. First, we will gain an understanding of how the trust that multiple human teammates have for their autonomous teammates can be tightly coupled, resulting in teammates' perceptions of trust affecting one another. Second, we will achieve an understanding of how trust, a complex human factor, can spread, propagate, and blend throughout multiple HATs who share a common goal and environment. Not only does the spread of trust between human teammates in HATs need to be understood, but for this research to be applicable, the interaction between multiple HATs should be considered. Unfortunately, it is difficult to perform such research that utilizes multiple HATs simultaneously. However, the teamwork experimental platforms available to the PIs, as well as their wealth of experience in human-autonomy teaming, allow for multi-HAT constellation research to now be conducted, serving as a landmark in the human-autonomy trust research domain. Third, we will consolidate the wealth of knowledge gained from mixed-methods experiments to provide an empirical tool specifically designed to measure the changes in one's trust of an autonomous teammate caused by the levels of trust their fellow team members or other teams express. Fourth, we will verify and quantify the effectiveness of common trust repair techniques with the goal of preventing the spread of distrust both within and between HATs. ☆

Dr. Christopher Myers, Senior Cognitive Scientist, 711 HPW/RHWM

Dr. Nathan McNeese, Principal Investigator, Clemson University

Dr. Jamie Gorman, Principal Investigator, Georgia Tech

DISTRIBUTED, HETEROGENEOUS TEAMING SOLUTIONS

TEAM KICKSTARTER:
APPROACHES TO FORM AND SUPPORT FLUID AND DISTRIBUTED TEAMS

Future Joint All-Domain Command and Control (JADC2) operations will require disparate and distributed team members to rapidly assemble, perform a mission, and then disassemble frequently with little-to-no previous experience with their teammates. While the flexibility associated with temporary team formation may be necessary to meet the future's dynamic battlefield, research suggests that teams with transient membership–sometimes referred to as *fluid teams*–often underperform compared to teams with stable memberships. This gap is partly due to teamwork factors that are necessarily influenced by team experience, such as team efficacy, transactive memory, and commitment to team goals (Bushe & Chu, 2011). However, fluid teams are common in some industries, such as aviation (flight crews) and healthcare (surgical teams), providing an opportunity for AFRL to leverage and extend previous research in these areas to optimize DoD teams.

This line of effort in the Collaborative Interfaces and Teaming Core Research Area is focused on mapping important factors and processes that influence teamwork in fluid and distributed teams, including swift trust and common ground. Our goal is to create technologies that support rapid formation of fluid teams and interventions that increase teamwork.

We have initiated two projects to support these goals. First, we are engaged in a systematic review of the relevant literature to characterize what has already been accomplished in this area and to identify gaps that need to be addressed. This review is being conducted in collaboration with Dr. James Driskell (Florida Maxima Corporation, Rollins College), a leading expert on teamwork. Second, we are planning a tech sprint to rapidly develop and test an interface for Airmen and Guardians to allow them to quickly establish common ground with distributed teammates before a mission. ☆

Dr. Michael Tolston, Research Psychologist, 711 HPW/RHWCT

Dr. Greg Funke, Senior Research Psychologist, 711 HPW/RHWCT

Graphic by Monch Publishing Group

SWIFT TRUST

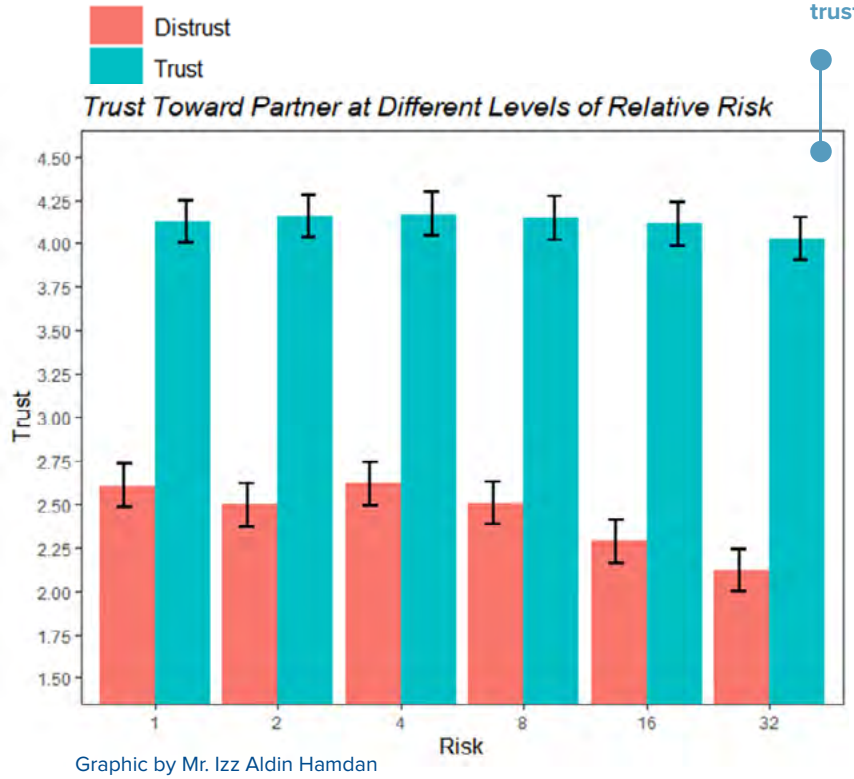
This program investigates trust formed quickly (i.e., swift trust) in ad hoc teams. At the end of FY 20, we published a cognitive task analysis (Capiola et al., 2020) investigating the antecedents to swift trust in multi-domain command and control (MDC2) contexts. In that analysis, eight antecedents emerged, with further analysis implying that swift trust is a relevant and emergent state in MDC2 that facilitates reliance. In FY 21, we designed experiments leveraging a custom testbed to quantify the influence of a contextual factor (i.e., differential risk) and a novel antecedent identified in our task analysis on shaping swift trust of experimental domains simulating aspects of joint all-domain command and control (JADC2) contexts. We adapted our testbed for online data collection and assessed an appropriately powered sample to investigate the effects of (ostensible) partner trustworthiness on trust-relevant criteria and how these effects are qualified by differential risk—a variable we instantiated for experimental research, considering risk is not uniform across nodes of the JADC2 enterprise. Preliminary analyses show several instances in which the influence of (un)trustworthy behavior in distributed ad hoc teams is moderated by the differential risk under which these ad hoc teammates operate. Soon, we aim to launch another data collection, investigating the role of partner self-awareness—a novel antecedent highlighted in our CTA—on the trust process in distributed, computer-mediated teams. ☆

Dr. August Capiola, Research Psychologist, 711 HPW/RHWCT

Screengrab by Mr. Dexter Johnson

Round results page explaining how a partner's trustworthiness and contextual risk affects one's score in the custom testbed.

Differential risk moderates the effects of partner trustworthiness on participants' trust toward their ad hoc teammate.



Graphic by Mr. Izz Aldin Hamdan

"In that [cognitive task] analysis, eight antecedents emerged, with further analysis implying that swift trust is a relevant and emergent state in MDC2 that facilitates reliance."

— Dr. August Capiola
Research Psychologist, 711 HPW/RHWCT

TEAM METRICS

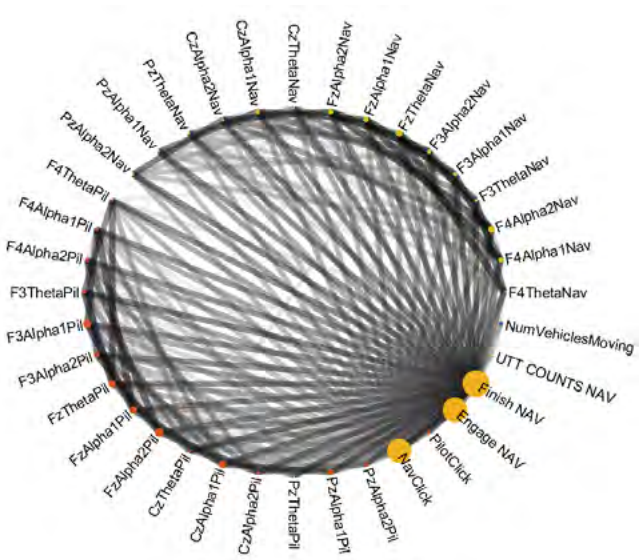
The Air and Space Forces rely on teams of Airmen and Guardians for many critical operations. These teams often face difficult situations, such as high-tempo operations, distributed team environments, long-shift durations, and decision making under uncertainty, that can lead to team under-performance and jeopardize mission success. Therefore, monitoring and evaluating teams is a critical AF endeavor. At AFRL, we are developing and applying cutting-edge approaches to uncover the drivers and consequences of successful team-level coordination. For instance, we have integrated measures of communication, team-level physiological and behavioral patterning, and environmental variables to create high-resolution, network-based models of team dynamics (see Figure). These methods can be used to identify patterns of development in teams during learning and maturation along with diagnosing why teams may be under-performing. Our goal is to advance our approaches into technologies that automatically identify and diagnose team states and dynamics in order to maximize the value of teaming outcomes. Highlights of recent research include development of a novel metric of machine agent influence on team decision making processes, and, in collaboration with Dr. Anne-Marie Brouwer (TNO), new measures of team-level synchronous auditory attention derived from brain activity. ☆

Dr. Michael Tolston, Research Psychologist, 711 HPW/RHWCT

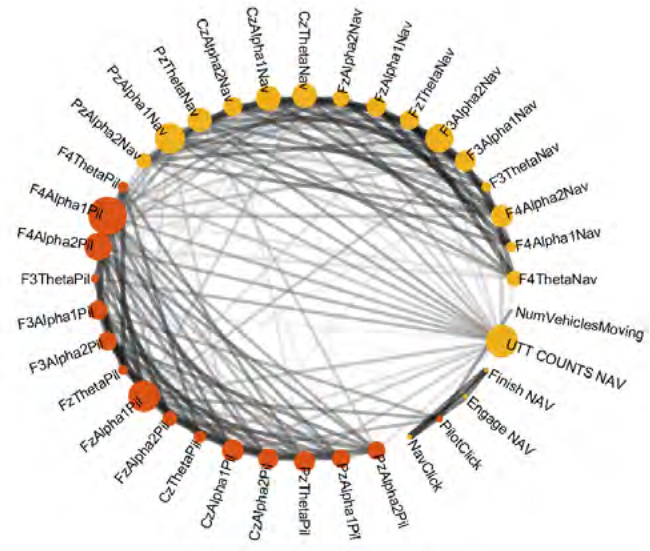
Dr. Greg Funke, Senior Research Psychologist, 711 HPW/RHWCT

Graphics by Dr. Michael Tolston

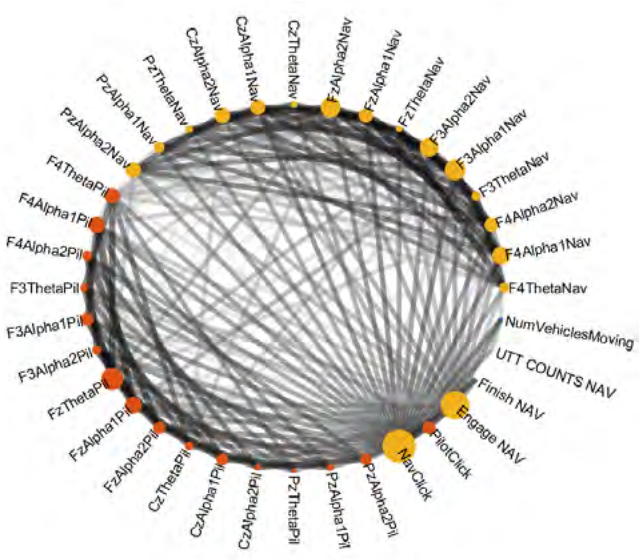
High Effective Teaming



Low Effective Teaming



Moderate Effective Teaming



JOINT ALL-DOMAIN OPERATIONS:
BATTLE MANAGER INTERFACES

This project demonstrated proof-of-concept user interfaces designed for a Joint All-Domain (JAD) Battle Manager to execute mission objectives of tracking and apprehending a high-value asset using cyber means. The Intelligent Multi-Vehicle Planner with Adaptive Collaborative/Control Technologies (IMPACT) testbed provided the baseline capability, including the Play Calling Tile with associated Play Workbooks, the Task Manager Tile, and the Play Sandbox interfaces. These user interfaces were expanded to include two new cyber plays–Cyber Locate and Cyber Inform. The Cyber Locate play was used to track the cell phone of the high-value asset, and the Cyber Inform play used social media accounts and posts to influence the high-value asset's movements. This project led to a follow-on effort that included the ability of an intelligent agent to reason and recommend cyber play to achieve a specific mission objective. Also, a proposed new start entitled JADPACT will continue to integrate Joint All Domain Battle Management user interface concepts and human-machine teaming solutions in IMPACT. ☆

Dr. Kristen K. Liggett, Principal Human Factors Engineer, 711 HPW/RHWCI

Dr. Gina Thomas, Research Psychologist, 711 HPW/RHWCI

Ms. Gloria Calhoun, Principal Human Factors Psychologist, 711 HPW/RHWCT

Ms. Jessica Bartik, Research Psychologist, 711 HPW/RHWCT

Dr Elizabeth Frost, Research Psychologist, 711 HPW RHWCT

Mr. Allen Rowe, Senior Computer Engineer, 711 HPW/RHWCI

Graphics by Ms. Arielle Stephenson

PLAYING "CYBER" IN JOINT ALL-DOMAIN WARGAMES:
UPDATING MERLIN TABLETOP WARGAMING MODULE

To be effective in Joint All-Domain (JAD) operations, the Air Force needs the capability of wargame JAD effects to better understand the entire range of courses of action and determine optimal integration of these effects. "Cyber" is rarely played in wargames and when it is, it is not always played realistically. For instance, there are assumptions that capabilities will automatically be available to use, they will be available in infinite quantities, and will always be 100% successful when they're used. Consideration must be given to the resources (time, manpower, finances, etc.) it takes to develop these effects if the wargame is designed to give us meaningful and useful results. AFRL's Wargaming Division contracted the Center for Naval Analyses to develop Merlin–a cyber module for tabletop wargames. Merlin was originally created in Excel but has recently been converted to Python, providing an opportunity for human factors engineers in RH to design a Merlin player interface. The interface is designed to guide game players in creating cyber "tradecraft" and to allow them to visualize the required resources for various tradecraft so that they can reason over trade-offs and make decisions about if, when, and how to effectively use cyber during game play. The next steps include expanding the use of the Merlin module from tabletop exercises to virtual-wargaming activities in the Advanced Framework for Simulation, Integration and Modeling (AFSIM) environment. ☆

Dr. Kristen K. Liggett, Principal Human Factors Engineer, 711 HPW/RHWCI

> **To be effective in Joint All-Domain (JAD) operations, the Air Force needs the capability of wargame JAD effects to better understand the entire range of courses of action and to determine optimal integration of these effects."**
>
> **— Dr. Kristen K. Liggett**
> Principal Human Factors Engineer, 711 HPW/RHWCI

MULTISENSORY PERCEPTION AND COMMUNICATION

Dr. Griffin Romigh
Multisensory Perception and Communication
Core Research Area Lead, 711 HPW/RHWS

Effective communication of information is critical for the success of operators in any domain and every Air and Space Force mission. However, such missions routinely take place in settings that are cognitively and environmentally complex, comprising multiple sources of stimulation, intense task demands, and situation uncertainty that pose unique challenges for effective portrayal of complex information. The Multisensory Perception and Communication Core Research Area identifies and exploits the underlying sensory and cognitive mechanisms that mediate perception and communication in order to inform the development of multimodal interfaces and advanced communication technologies. The Core Research Area prioritizes two lines of effort:

1. Exploiting Perceptual Abilities (EPA)
2. Enhancing Operational Communications (EOC)

Under EPA, research focuses on identifying the mechanisms and phenomena associated with sensory encoding, crossmodal interactions, and multisensory integration in order to inform the development of advanced multimodal interfaces. In addition, research is conducted to examine the inherent constraints on perceptual and cognitive processing that limit performance in situations requiring the monitoring and management of multiple information sources. This knowledge is being exploited to enhance, or inhibit, task performance. Finally, advanced methodologies and technologies are being employed to recreate many of these phenomena virtually to support distributed operations.

The objective of EOC is understanding and characterizing the underlying processes that support effective human-human communication in order to inform the development of natural human-machine communication interfaces. Research focuses on the identification of acoustic, linguistic, and task strategies humans employ to effectively transmit spoken information to achieve mutual understanding and avoid miscommunication. This research informs the development of synthetic agents that can monitor task environments and intelligently time the presentation of information support. Finally, advanced systems that are made to manage complex communication situations are being developed, tested, and transitioned to the Warfighter.

Both EPA and EOC conduct research in unique laboratory facilities designed for rigorous stimulus control; these same stimuli are also examined in laboratories that recreate the

THE TWO LINES OF EFFORTS

Exploiting Perceptual Abilities (EPA)

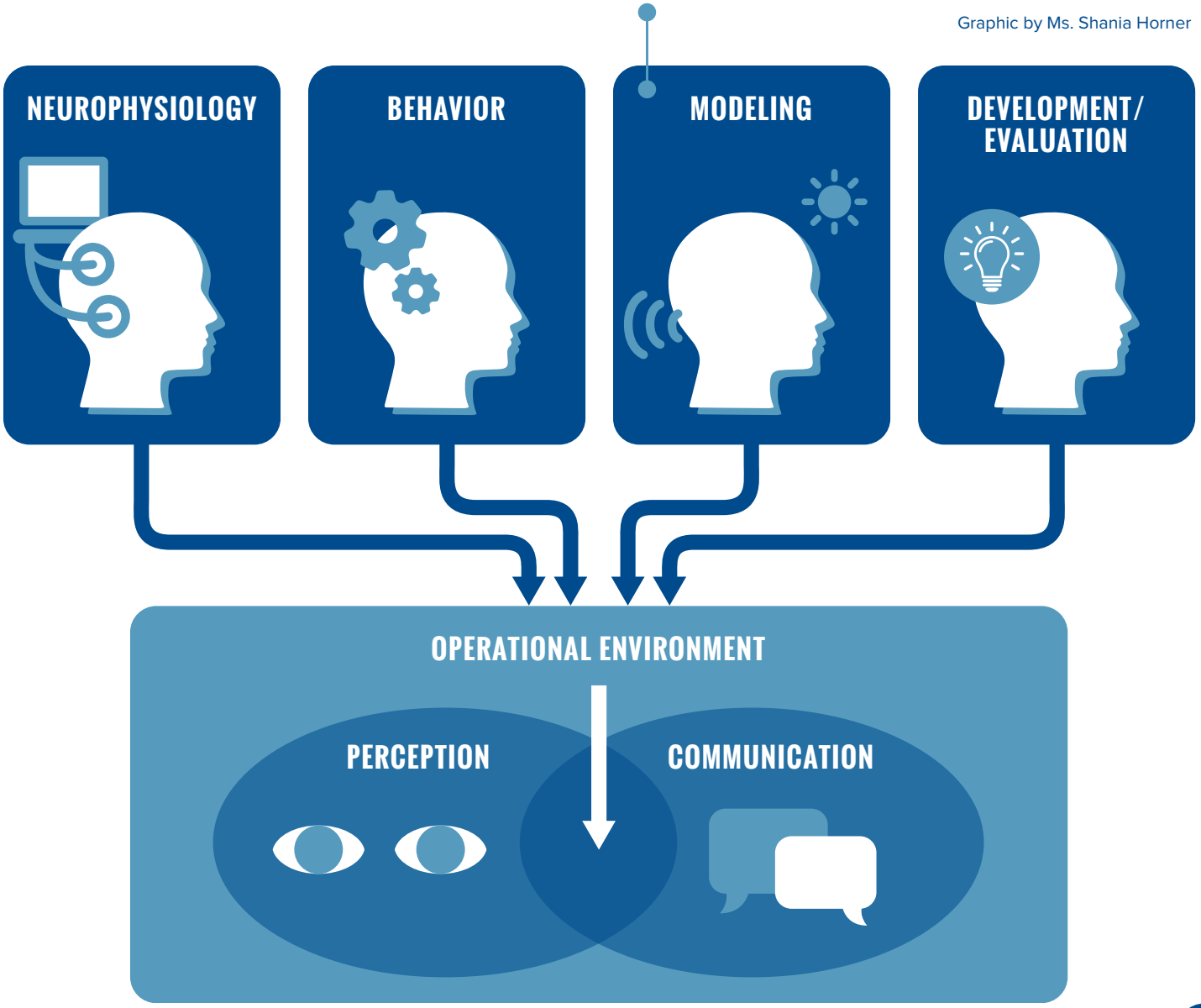
Enhancing Operational Communications (EOC)

stressors inherent in operational settings. Both performance-based behavioral and neurophysiological data are collected to identify neural correlations of perception and inform computational models that can guide multimodal display development. This approach ensures that technologies created by this CRA will support the presentation of the right information at the right time and in the right way. ☆

Dr. Griffin Romigh, Multisensory Perception and Communication Core Research Area Lead, 711 HPW/RHWS

Below represents the scientific approach of the Multisensory Perception and Communication team.

Graphic by Ms. Shania Horner



EXPLOITING PERCEPTUAL ABILITIES (EPA)

MANAGING MULTISENSORY INFORMATION THROUGHPUT

In our everyday interactions with the world, individuals are routinely overwhelmed with multiple, concurrent information streams from many sources, which they must effectively process and integrate. This is particularly true in military operational environments, where stimuli and task complexities are especially high and time constraints are severe. However, because operators have only a finite amount of attentional resources to filter and process this information, their vulnerability to sensory overload and decision errors often increases. In novel scenarios, operators must filter the information necessary to complete their assigned tasks while monitoring their environment and the actions of their teammates to maintain situational awareness and quickly intervene in response to unexpected events.

The Managing Multisensory Information Throughput task uses neurophysiological and behavioral measures to examine the strengths and limitations of human perception and cognition while defining principles that allow for the proper integration and distribution of multisensory data in operational environments. We design displays that tailor to individual operators, taking into account both capabilities and proclivities of individuals, to ensure that information is processed efficiently, while overall decision making and performance are optimized. To best do this, we look at how sensory information processes interact, starting from basic neurological and psychophysical principles; we establish the end goal of creating a user-centered suite that dynamically recommends the optimal multisensory-information display for the person, mission goals, technology, and physical constraints of the environment.

Basic research includes the assessment of integration and multisensory facilitation (i.e., the effect of one sense increasing the effectiveness of another) to improve performance and reduce the risk of perceptual-central tunneling personalization. Additionally, research entails the use of statistical tools, such as General Recognition Theory, that enable us to strategically distribute multisensory information in a customizable display using mathematical performance-based predictions, neurobehavioral correlation. Also, the research includes the identification of neurophysiological markers and behavioral correlations of multisensory perception, attention, decision processes, and distribution. Our research also includes assessment of attentional processing capacity and performance tradeoffs to suggest when and how to off-load information from one sense to another within and between tasks to effectively mitigate perceptual and cognitive overload. Our framework will unify and personalize the sensory experience to inform display design, optimize users' multitasking strategies, inform interface adaptation, and supply pertinent feedback to users and teammates so that they can work together and perform their jobs more effectively. ☆

Dr. Elizabeth Fox, Research Psychologist, 711 HPW/RHWS

Dr. Paul Havig, Senior Research Psychologist, 711 HPW/RHWS

Graphic by Mr. Will Graver

INFLUENCING PERCEPTION

Humans have a remarkable ability to parse complex visual and auditory environments, forming perceptual objects to extract relevant information. However, this process is imperfect and often incomplete, leading to missed critical information. Understanding the abilities and constraints of human perception and cognition informs the development of advanced capabilities that will enable the US Air and Space Forces to influence the human information processing space. The Influencing Perception task leverages a long history of research at AFRL, examining the limits of perception and communication to better understand where the failures in perceptual analysis occur in real-word environments and how to best control the nature of those failures. We are identifying new, optimal approaches for describing various aspects of complex, real-world scenes (presence/absence of critical objects, spatial properties of the objects in the scene, and overall scene dynamics) through advanced, statistical techniques and computational models developed with an understanding of human perception. These descriptions are grounded in measurements of real-world auditory and visual environments and leverage a vast database, describing dynamic military assets. Small Unmanned Aircraft Systems (SUAS) are of particular interest due to their rising popularity of use in military missions. Synthetic representations of real-world scenes have been generated for greater laboratory control in order to bind the development of perceptual models, which will subsequently be tested in those same real-world environments. A human detection performance model is being developed based on the results of human behavioral studies. Human neurophysiological studies are also being conducted to study evidence accumulation to gain an understanding of the neural process for aircraft detection. We will create situations that both facilitate and disrupt this analysis process. ☆

Ms. Hilary Gallagher, Research Engineer, 711 HPW/RHWS

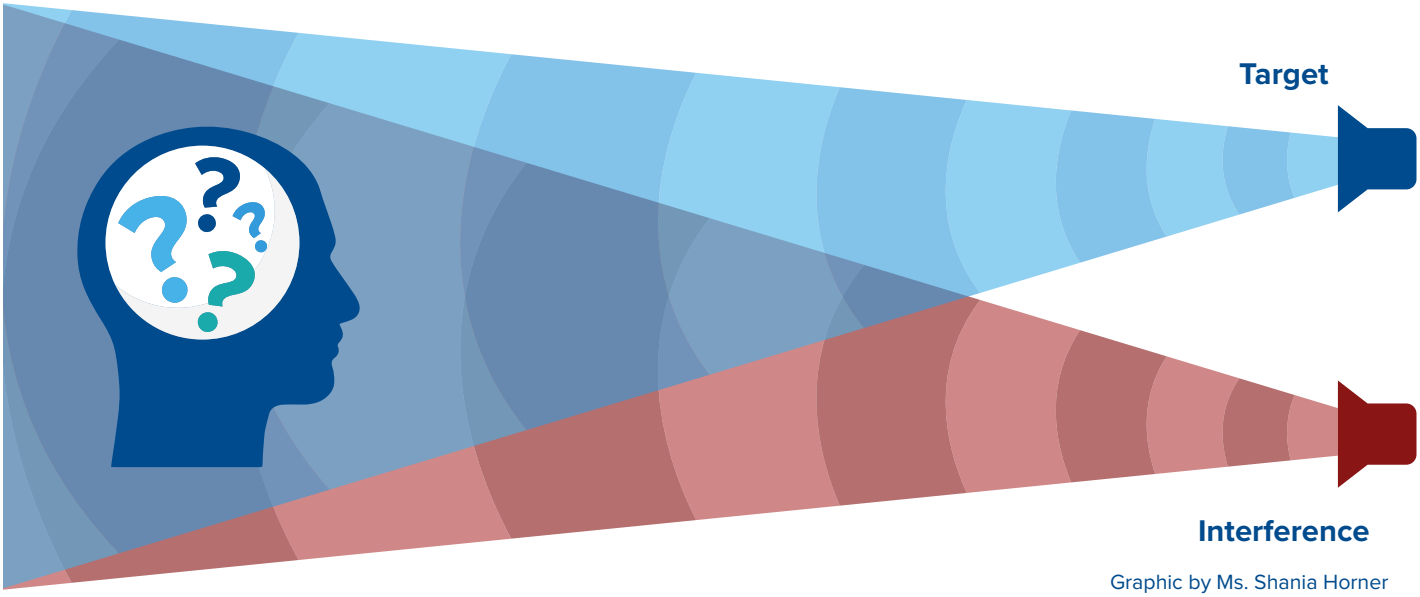
Dr. Eric Thompson, Research Engineer, 711 HPW/RHWS

HOW?

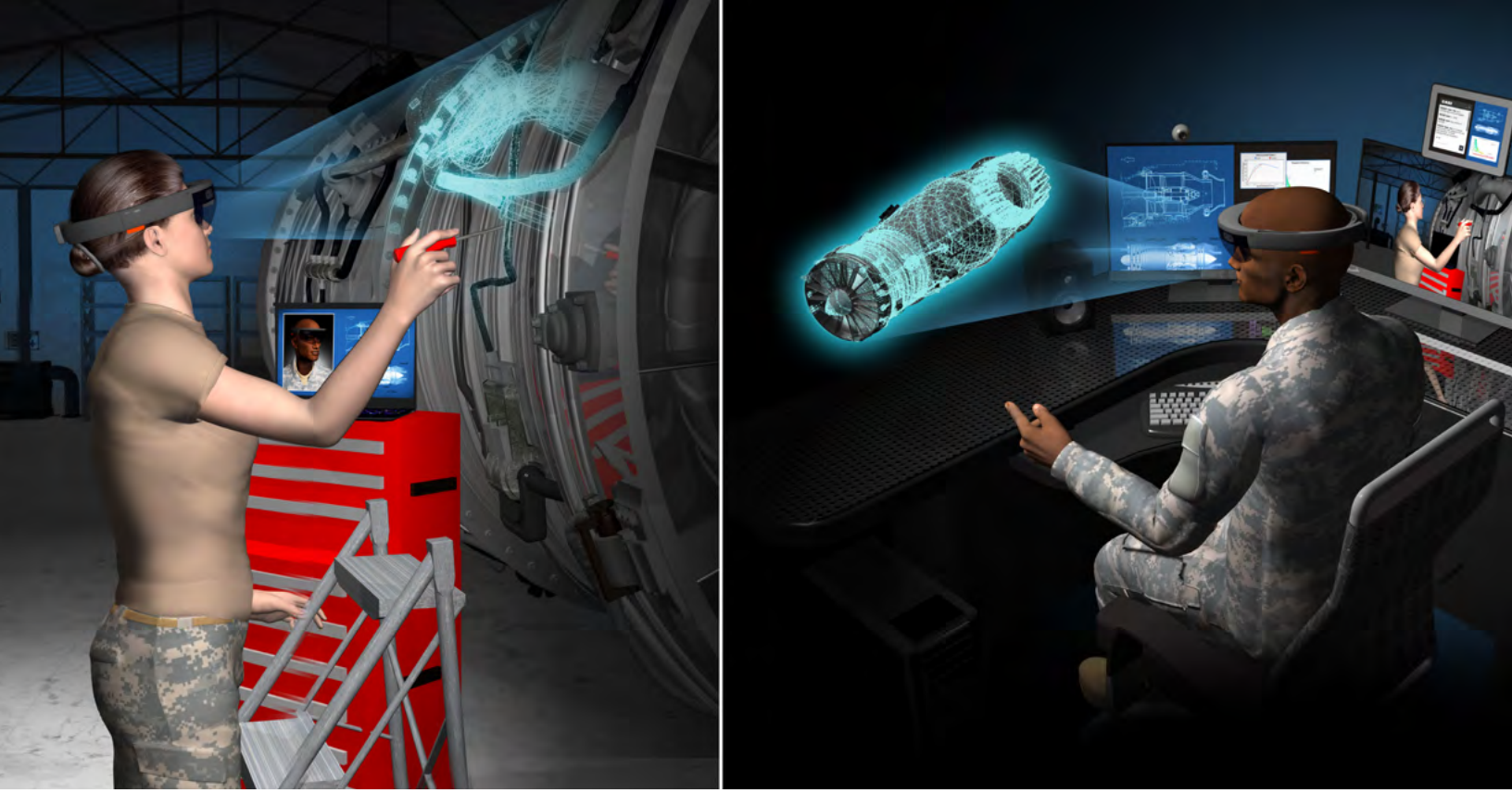
Through measurements of both human behavior and neural activity

WHY?

To identify markers of perceptual processes that reflect successful analysis



Graphic by Ms. Shania Horner



Graphic by Mr. Will Graver

IMMERSIVE VIEWPOINT SHARING

Future Air and Space Force operations will continue to grow in complexity, requiring coordinated activities across distributed agents, supporting missions taking place across multiple operational domains. This future will force significant change for the concept of mutual support. Affording flexibility of command and control across remote distances and a variation of the right effects at the right time can be a critical warfighting capability.

The goal of the Immersive Viewpoint Sharing research task is to develop and evaluate immersive interfaces, employing virtual-, augmented-, and mixed-reality technologies to enable operators a more intuitive understanding of remotely-sensed situation information. This will be informed by leveraging state-of-the-art hardware and interface design solutions to determine interactions between display capability and perceptual constraints. In our scenario, a human operator is the occupant of a "piloted" vehicle while simultaneously controlling other vehicle systems used as forward sensor platforms for purposes of tactical-situation awareness and weapons employment. The human operator will utilize a helmet-mounted display (HMD) approach to enable augmented-, virtual-, and mixed-reality multisensory appreciation of a selected remote vehicle environment from the perspective of the remote sensor. This interface approach gives the human operator the ability to 'seat hop' and temporarily control and experience the actions of the remote vehicle firsthand while ownership is maintained via automation. Our 'art of the possible' approach toward design and evaluation will begin by utilizing the functionality of a high-resolution, virtual-reality environment to mock-up a prototype interface configuration for experimentation. This will be performed in our Pilot/Vehicle Interface research facility equipped with multiple reconfigurable fight simulation stations. A number of research questions are being explored as a part of this effort, including:

- What is the nature of the information that should be provided?
- What level of multisensory fidelity is required to support effective shared experiences?
- How do measures of presence and immersion vary as a function of the level of recreated information?
- To what degree do these constructs correlate with performance enhancements? ☆

Dr. Eric Geiselman, Senior Engineering Research Psychologist, 711 HPW/RHWS

Dr. Paul Havig, Senior Research Psychologist, 711 HPW/RHWS

PERCEPTUAL ENVIRONMENT MODELING AND SIMULATION

Our lives are filled with acoustic stimuli that continuously provide information about objects and events in our environment. Indeed, our experiences are shaped by the sounds around us, which provide context, connect us to our world, and yield information to ground our decision-making processes. Military environments are acoustically rich and complex and can play a major role in the outcome of a mission. On the one hand, the extremes of the acoustic battlespace can impact communication effectiveness (including both speech intelligibility and the interpretation of warning signals) and can have a deleterious effect on advanced technologies, such as automatic speech recognition to function effectively. Conversely, the acoustic environment may be exploited by operators, consciously or subconsciously, to acquire information about the status of systems with which they are interacting, where a sudden change in the acoustics may signify an emergency (e.g., wind over a cockpit canopy indicating an unintended change in attitude or airspeed, or an impulsive 'pop' signaling a tire blow-out). When simulating operational environments for training, interface development, and teleoperation, it is critical to provide a genuine representation of the acoustic environment. To 'train the way we fight' (among other requirements), the Perceptual Environment Modeling and Simulation task builds physical acoustic models of noise environments based on first-principle state variables as determined from real-world measurements. These models are being used to recreate acoustic environments at the location of the operator and will be integrated into simulation architectures such that these environments change in realistic ways (e.g., with changes in airspeed, attitude, or air conditioning fan speed). To enhance realism and perceptual accuracy, communication signals will be processed so that they emulate the degradation that may occur through military communication systems. ☆

Dr. Frank Mobley, Research Physicist, 711 HPW/RHWS

Dr. Alan Wall, Research Physicist, 711 HPW/RHWS

"In effort to... 'train the way we fight', the Perceptual Environment Modeling and Simulation line of effort builds physical acoustic models of noise environments based on first-principle state variables as determined from real-world measurements."

— Dr. Frank Mobley
Research Physicist, 711 HPW/RHWS

Graphic by Mr. Frank Mobley

ENHANCING OPERATIONAL COMMUNICATIONS (EOC)

COMMUNICATION RESILIENCY

Human-machine teaming aims to meld human cognitive strengths with the unique capabilities of smart machines to create intelligent, resilient teams. A key challenge within human-machine teaming is the establishment of natural human-machine interfaces that enable effective communication and coordination. In particular, interfaces based on spoken language, central to natural communication in human-human teams, hold great promise as conduits for efficient information flow in human-machine teams. Research conducted under the Communication Resiliency task focuses on developing an understanding of communication processes in human-human teams–specifically, how humans repair communication errors and establish a shared understanding (common ground) in dialogue. The results from this research inform the development of human-machine spoken dialogue systems with the capability to quickly recover from communication difficulties and reduce the deleterious effects of miscommunication on team performance in collaborative tasks. The research is grounded in theoretical models developed to describe natural communication in "everyday" conversation but seeks to modify/extend these models to operationally-relevant military communication environments, which tend to be more extreme and are characterized by greater perceptual, cognitive, and environmental complexity (e.g., high-tempo operations in noisy air operation centers). Such environments increase the likelihood of communication uncertainty, ambiguity, and error. In addition, idiosyncrasies associated with military communication protocols may pose additional challenges to existing models. The approach taken in this task involves the development of a number of laboratory-based collaborative tasks that:

- Capture the dynamics of real operational communication environments.
- Use experimental paradigms to simulate conversational repairs and grounding in voice-user interfaces.
- Develop spoken and natural language processing models capable of creating conversational alignment and recovery.
- Evaluate the overall human-machine team performance within systems that employ these capabilities.

The long-term objective of this research effort is to provide insight into the flexible conversational strategies that humans use to circumvent communication challenges. These challenges typically occur when humans inform the development of natural communication interfaces that inevitably encounter similar communication errors, specifically within complex and time-critical human-machine interactions. ☆

Dr. Sarah Bibyk, Research Psychologist, 711HPW/RHWM

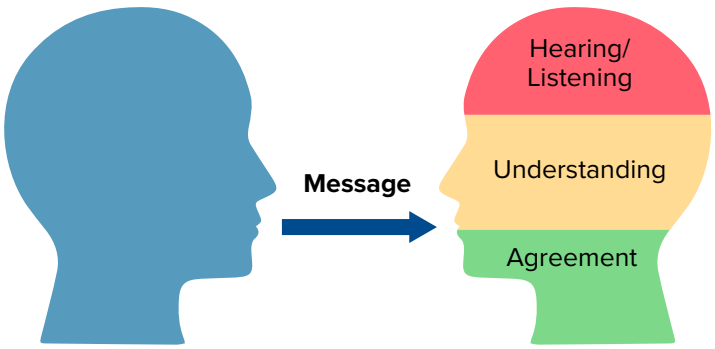
RESEARCH IS FOCUSED ON:

Developing an understanding of communication processes in human-human teams

WHY?

To inform the development of human-machine spoken dialogue systems with the capability to quickly recover from communication difficulties

To reduce the deleterious effects of miscommunication on team performance in collaborative tasks



Graphic by Ms. Shania Horner

CONTEXT-AWARE COMMUNICATION INTERFACES

As the warfighting mission becomes more complex, operators must manage more information from disparate sources, attend to a greater number of concurrent tasks, and effectively collaborate with distributed team members across domains and mission sets. These military missions are likely to have time-critical components that rely on the expeditious presentation of up-to-date information to inform ongoing tasks. Therefore, the possibility that new sources of incoming information will interrupt ongoing task flow is increased, resulting in a greater potential for human error and, consequently, reduced mission effectiveness. To this end, the goal of the Context-Aware Communication Interfaces task is to develop natural human-machine communication interfaces that effectively determine "what, when, and how" an intelligent agent should disseminate information to human team members that:

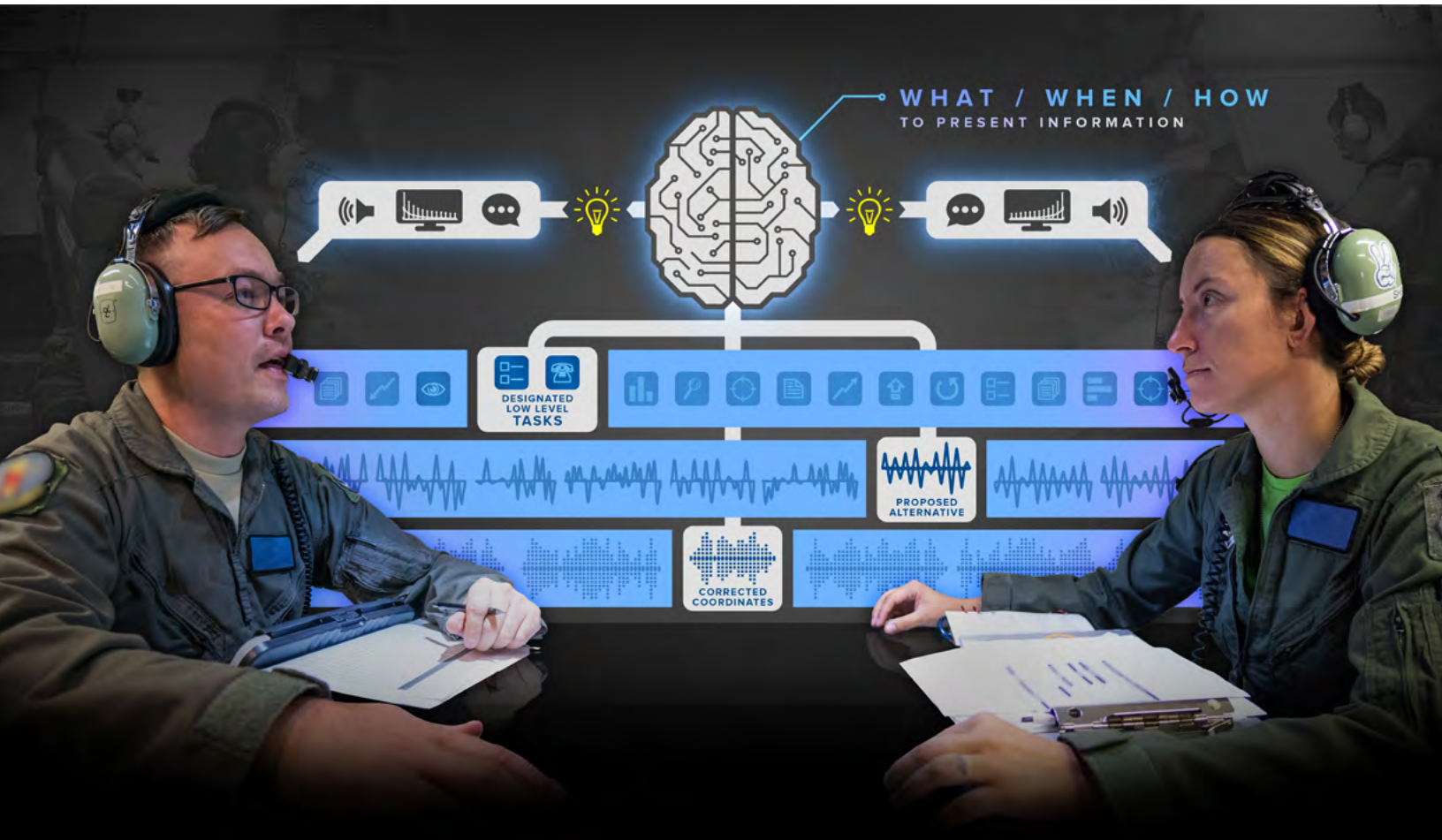
a) Adds to an operator's knowledge.

b) Is least disruptive to ongoing tasks.

Context-aware systems have the potential to automate low-level tasks and make inferences from ongoing, complex interactions, enabling operators to focus their efforts on critical, mission-level task objectives. Additionally, the complexity associated with real-operational environments undoubtedly leads to a rich set of cues that a system can leverage to determine the best time and manner to inform new sources of information and present this to teammates. Ongoing research and development attempts to reveal these cues while involving the testing and refinement of prototype intelligent interruption systems in complex, team-based task environments that capture the interactive nature of real-world military operations. Results from this research help reveal, specifically, how ill-designed information dissemination mechanisms impact overall decision quality, task-related response times, and the processes supporting task resumption. Research devoted to context-aware communication interfaces can inform specific modifications to intelligent systems within human-machine teaming in order to mitigate such deleterious effects on human performance. ☆

Dr. Griffin Romigh, Multisensory Perception and Communication Core Research Area Lead, 711 HPW/RHWS

Graphic by Mr. Will Graver



COMMUNICATION SCENE ANALYSIS AND DISRUPTION

Verbal communication is a vital piece of the success of every USAF mission, but significant challenges to effective communication exist, including interference from high levels of environmental noise, reduced fidelity from communication system distortion, and the need for operators to manage multiple simultaneous streams of information. The Communication Scene Analysis and Disruption task conducts research to quantify the impact of these barriers to successful communication. Of particular interest here is the degradation in intelligibility that can occur when a speech signal of critical interest is presented in an auditory scene that comprises additional stimuli that are similar along some perceptual dimension. This perceptual masking, or informational masking (IM), seems to occur because of a failure in the system's ability to effectively analyze an acoustic scene to segregate target from masker.

Research conducted under this task focuses on identifying these specific perceptual dimensions that lead to IM; the strategies humans may use to overcome its effects; and the quantification of the impact on overall human performance, including both task performance and listening effort. Importantly, understanding the specific information that listeners utilize to parse complex scenes has led to the construction and testing of specific stimuli in order to systematically manipulate IM, as well as the development of communication-enabling technologies that can make it easier to understand speech in these complex environments. Moreover, the ability to synthesize specific speech features in these complex scenes can be used to generate seemingly intelligible, but potentially ambiguous, speech signals that similarly impact overall task performance. These synthetic stimuli can be used in training environments to improve operator performance on complex listening tasks, as well as in the development of novel communication technologies. ☆

Dr. Eric Thompson, Research Engineer, 711 HPW/RHWS

INTERACTIVE COMMUNICATION MANAGEMENT

The goal of interactive communication management research is to increase an operator's ability to access and utilize verbal information within the fight. Missions rely on the effective communication between teammates, and our operators serve in cognitively demanding roles where it's critical to have communication interfaces that overcome informational overload and the loss of situation awareness due to attentional tunneling. Our research combines advancements in our understanding of human perception with recent developments in AI and machine learning to overcome limitations of classic communication channels. This includes limitations like the perishable nature of standard voice-based communications, the visually demanding nature of text and chat, and the cross-modal conflicts that can occur when attempting to process both simultaneously. By utilizing state-of-the-art advancements in automatic speech recognition and natural language processing, we are developing techniques to enable intuitive and flexible interactions with the communication channels themselves. This will allow operators to dynamically remove non-critical communication traffic and receive the most relevant information while in the modalities best suited for their mission context. Current efforts include increasing the flexibility of and enabling natural interaction with previously-developed communication management technologies, as well as developing advanced communication technology suitable for stand-alone, resource-limited mobile platforms. ☆

In today's world, there are so many ways to communicate. We focus on combining advancements in our understanding, in addition to developments in AI and machine learning. This enables enhanced interaction in communication.

Dr. Griffin Romigh, Multisensory Perception and Communication Core Research Area Lead, 711 HPW/RHWS

THE GOAL OF INTERACTIVE COMMUNICATION MANAGEMENT RESEARCH IS TO INCREASE AN OPERATOR'S ABILITY TO ACCESS AND UTILIZE VERBAL INFORMATION WITHIN THE FIGHT."

— Dr. Griffin Romigh
Multisensory Perception and Communication
Core Research Area Lead, 711 HPW/RHWS

Graphic by Ms. Shania Horner

Published quarterly since 2001, Fight's ON! continues to serve as the Division publication for our partners and features innovative science and technology that is accelerating and revolutionizing readiness.

Fight's ON! Point-of-Contact
Patricia D. Wood, 711 HPW/RHW
patricia.wood.2@us.af.mil

Cleared / Case # AFRL-2021-3518

711 HPW/RHW Branches:

- Cognitive Models Core Research (RHWM)
- Personalized Learning and Readiness Sciences (RHWL)
- System Analytics (RHWA)
- Collaborative Interfaces and Teaming (RHWC)
- Multisensory Perception and Communication (RHWS)