AFRL FIGHT'S ON!

THE AIR FORCE RESEARCH LABORATORY

ISSUE 67
FALL 2022

WARFIGHTER INTERACTIONS AND READINESS DIVISION

I/ITSEC Edition

TABLE OF CONTENTS

AS AFRL CONTINUES ITS JOURNEY, OUR COMMITMENT TO THE WARFIGHTER IS STRONGER THAN EVER. WE REMAIN FOCUSED ON THE FUTURE AND READY TO TAKE ON OPPORTUNITIES THAT LEAD TO DISCOVERIES AND SOLUTIONS FOR FUTURE OPERATIONAL DOMINANCE."

— Major General Heather L. Pringle
Commander, Air Force Research Laboratory

Photo by Airman 1st Class Tiffany Del Oso

Graphic Designer **Mr. Will Graver**

Editor in Chief **Mr. Dave Hubbell**

Photo by J.M. Eddins Jr.

RHW

WARFIGHTER INTERACTIONS AND READINESS DIVISION

U.S. AIR FORCE

The United States Air Force stands at the forefront of progress, and ongoing initiatives reveal the promising tomorrow that can - and will - be earned. The Air Force Research Laboratory (AFRL) improves the preparation and execution needed for successful military operations through the Warfighter Interactions & Readiness Division (RHW). The Division is committed to establishing and securing the warfighter's legacy of success and future efficiency by working with other Divisions and researchers to produce findings and 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.* ☆

DR. LOUISE CARTER

Division Chief,
711 HPW/RHW

"Our Airmen and Guardians need to be prepared for the mission, aided by autonomous technologies, provided information for smart decision making, and have the tools to act quickly in the face of changing threats."

— Dr. Louise Carter
Division Chief, 711 HPW/RHW

RHW

WARFIGHTER INTERACTIONS AND READINESS DIVISION

Our RHW team members have been participating in a variety of planning activities addressing the Secretary of the Air Force's Operational Imperatives. It is clear that our research to enable and enhance warfighter cognition and decision making, is critical across the OIs. Our Airmen and Guardians need to be prepared for the mission, aided by autonomous technologies, provided information for smart decision making, and have the tools to act quickly in the face of changing threats. We have focused our research portfolio on the cognitive aspect of the mission, specifically enabling our warfighters at every level for decision superiority. Joint All-Domain Operations is a lens for much of our research. For example, the Agile Warfighter Interfaces CTC has early research in distributed teaming, team collaboration, and interfaces for handling multi-domain assets. The Training CTC is building on our long history of simulation and training to investigate what the future of synthetic operational testing and training should look like.

Our success in integrating our Adaptive Teamwork with Layered Airman-Machine-Interfaces & Systems (ATLAS) with the Skyborg program has led to multiple requests for our expertise and technologies. We are participating in multiple cross-technical directorate efforts where AFRL hardware developers recognize the need for smart human-machine teaming. Mission planning, space domain awareness, and intelligence analysis are just some of the operations that our researchers are teaming up to make a difference. RHW is discovering and developing human-machine technologies to ensure weapon system effectiveness. We've learned that integration and teaming is our way of business. I'm excited for the future as we have big goals and a great team helping us achieve them. ☆

Dr. Louise Carter, Division Chief, 711 HPW/RHW

RHW ORGANIZATION

WARFIGHTER INTERACTIONS & READINESS DIVISION
711 HPW/RHW

DIVISION CHIEF
Dr. Louise Carter

DEPUTY CHIEF
Col Alfredo Rivera

BRANCH STRUCTURE

TECHNICAL ORGANIZATION

PRODUCT LINES

READINESS PL
Dr. Winston "Wink" Bennett

ANALYTIC TOOLS PL
Mr. Greg Beister

AIRMAN-MACHINE INTEGRATION PL
Dr. Deirdre Mahle

BRANCHES

RHWM
COGNITIVE MODELS
Lt Col Scott Storm

RHWL
CONTINUOUS LEARNING
Dr. John Camp

RHWS
SENSORY SYSTEMS
Ms. Jennifer Brown

RHWA
MISSION ANALYTICS
Dr. William Murdock

RHWC
COLLABORATIVE INTERFACES AND TEAMING
Dr. Terry Stanard

CORE TECHNICAL COMPETENCIES (CTC)

TRAINING CTC
Dr. Glenn Gunzelmann

ADAPTIVE WARFIGHTER INTERFACES CTC
Dr. Mark Draper

CORE RESEARCH AREAS (CRA)

COGNITIVE MODELS CRA
Dr. Megan Morris

PERSONALIZED LEARNING & READINESS SCIENCES CRA
Ms. Jennifer Winner

MULTISENSORY PERCEPTION & COMMUNICATION CRA
Dr. Griffin Romigh

SYSTEMS ANALYTICS CRA
Dr. Vincent Schmidt

COLLABORATIVE INTERFACES & TEAMING CRA
Dr. Chris Brill

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

READINESS PRODUCT LINE

The Readiness Product Line continues to support Science and Technology (S&T) innovation inside the Air Force Research Laboratory (AFRL), in the United States ops communities, and with our international partners. We're excited to continue to develop these advances, collaborations, and transitions. Other efforts this year include the FedLabs Open Mission Systems as well as efforts with the Joint Simulation Environment in the USAF, the multinational Distributed Synthetic Training initiative, and advances in the USAF's Simulator Common Architecture Requirements (SCARS) implementation.

We've also continued our augmented reality/virtual reality (AR/VR) content repository work and are wrapping up a two-year effort to culminate the peer-reviewed published AR/VR effectiveness literature and create summaries and a how-to guide to help future development and field evaluations. We're now collaborating with Dynepic to share this work inside the MOTAR® ecosystem. We're excited for this partnership to grow!

This year, we continued innovations in content and data standards supporting the Combat Air Forces migration to proficiency based training (PBT). Our AFRL core funded effort is supporting several operational toolset and data lake deployments, and Air Combat Command (ACC) is supporting collaborations with CAE and Plexsys to integrate our PBT tools within their mission training center baselines and to demonstrate and de-risk this integration for ops.

Continued on next page

DISTRIBUTED MISSION OPERATIONS NETWORK

Synthetic Operational Test and Training Infrastructure (S-OTTI)

Graphic by Mr. Will Graver

Secure connection to sites throughout the Air Force and around the world

Graphic by Mr. Will Graver

We have digitized over 20 years of paper-based instrumentation and created a construct-oriented taxonomy of those instruments and the associated data from them. This has allowed us to create some incredibly robust training effectiveness and performance measure data files we want to share out to AI and Machine Learning researchers to develop novel analytics and outcomes to better support ops decision makers.

We have received initial funding for a new, 4-year cross product line effort called Just In Time Multimission Airmen/Warfighters (JITMMA/W), funding to support S&T in Synthetic Operational Test and Training Infrastructure (S-OTTI), as well as outyear funding to extend this initial work to Joint All-Domain Ops.

The Gaming Research Integration for Learning Lab® (GRILL®) continued our strong support for USAF ops training and science, technology, engineering, and math (STEM) and we're continuing to pursue the exciting opportunity to establish a GRILL® at the US Air Force Academy (USAFA). Several of our recent activities are being showcased in the USAF booth and in the STEM exhibition.

You can read more about some of these efforts in this issue. Finally, we continue to grow our industry collaborations with several new Small Business Innovation Research (SBIR) efforts totaling over $6M in new activities. It's been a busy year. Stay tuned for what's next! ☆

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

NOT SO GRAND CHALLENGE

The Not So Grand Challenge (NSGC) is an ongoing research activity intended to enhance fighter pilot training by developing red force cognitive models to make smarter, more realistic adversary threats. Through a partnership of companies working across the Department of Defense (DoD) space (including Aptima, Charles River Analytics, CHI Systems, Discovery Machine, Eduworks, Soar Technology, Stottler-Henke, Tier 1 Performance Solutions, and Plexsys), the project team is developing a variety of cognitive models, a Digital Librarian to store and recommend the utility of these models, and a virtual Testbed to access and engage with these models in a distributed work environment.

COGNITIVE MODELS

Engagement in exercises is a proven way to enhance pilot training, particularly when conducted against experts in tactical flight. The use of computer-generated forces (CGFs) allow for readily accessible and adaptive components to engage with pilots in a cost effective way. However, a gap exists in the development of intelligent adversary entities that go beyond the predictable behavior of red forces that presents the potential for negative transfer of bad lessons to real engagements. The NSGC project applies multiple approaches to developing intelligent adversary agents. Some apply a machine learning technique to learn scenario states and associated pilot actions. Others develop artificial intelligence (AI) pilot agents by generating automated measurements (i.e., data) of agent performance in scenario runs, and encoding subject matter expert-guided (SME-guided) interpretation of those data into their diverse cognitive modeling architectures.

NSCG MILESTONES

Completion of Testbed and Client-side software to enable cognitive model development

Completion of models reflecting 1v1 pilot performance in a series of basic scenarios

Completion of models reflecting 2v2 pilot performance with more complex scenarios

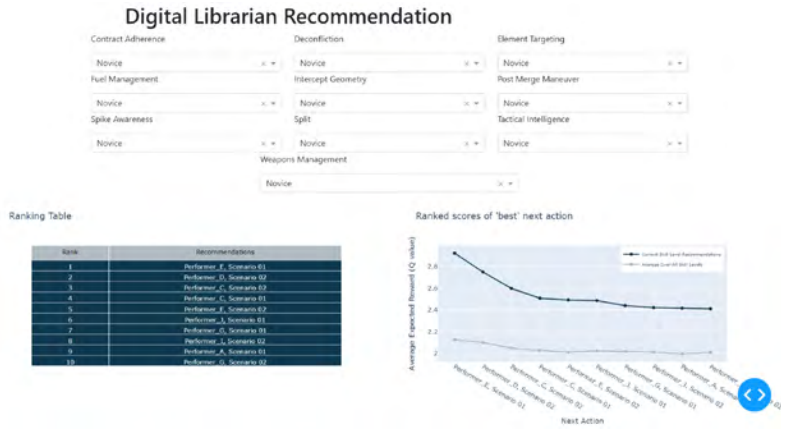
Continued on next page

NOT SO GRAND CHALLENGE (CONTINUED)

Graphics by Mr. Will Dupree and Mr. Luke Waggenspack, Aptima

DIGITAL LIBRARIAN

The Digital Librarian is a machine learning (ML) tool developed to aid in the training of pilots who will interact with the Testbed in the future. The end goal is to enable pilots being trained, and the personnel overseeing the training, to intelligently select available scenarios and artificial intelligent (AI) agents in a way that maximizes the pilot's acquisition and building of skills. As the Testbed grows in complexity with a larger number of AI agents and possible scenarios to train against, the challenge of choosing the best training environments to maximize a set of skills becomes intractable across high dimensional tracking metrics for any single human to perform. In this case of high dimensionality, the Digital Librarian will leverage information about current skill level, scenarios, and AI agents to provide a ranked list of recommendations it perceives as the best scenario and agent pair for a human pilot to train against with the goal of increased effectiveness and skill growth. The Digital Librarian uses Markov Decision Process (MDP) and reinforcement learning (RL) inspired algorithms to optimize what is the next best "sparring scenario" for a given pilot. In the top right image we share a prototype of the Digital Librarian. It is an interactive dashboard (currently using simulated stand-in data) that allows a user to set a trainee's current skill level and outputs recommendations of best-next-actions given the data it learned from. The left side table gives rank order of scenario/agent combinations. The right figure shows trend lines of expected outcome reward (larger is better) for current skill level in blue and average reward across all skill level combinations in gray. Reward here is the numerical value assigned to expected outcome of growing in skill level.

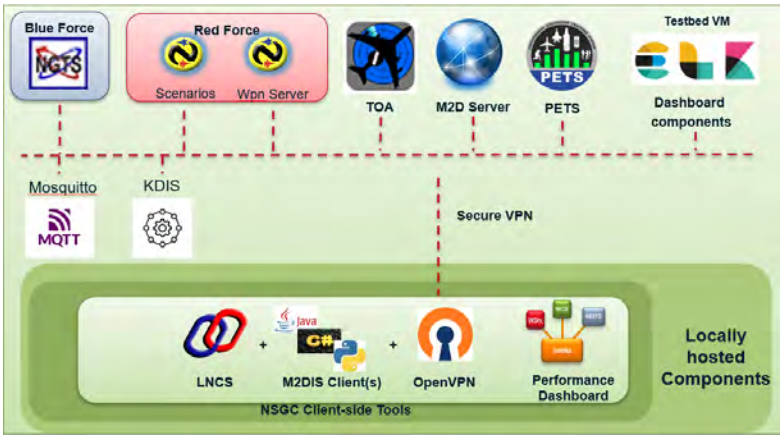


NSCG FUTURE WORK

Demonstrate model capability to perform numerous tasks: air refueling, combat air patrol maneuvering, bomber escort, and no fly zone avoidance

Develop Digital Librarian to ingest pilot performance measurements to determine which model is best for specific scenarios

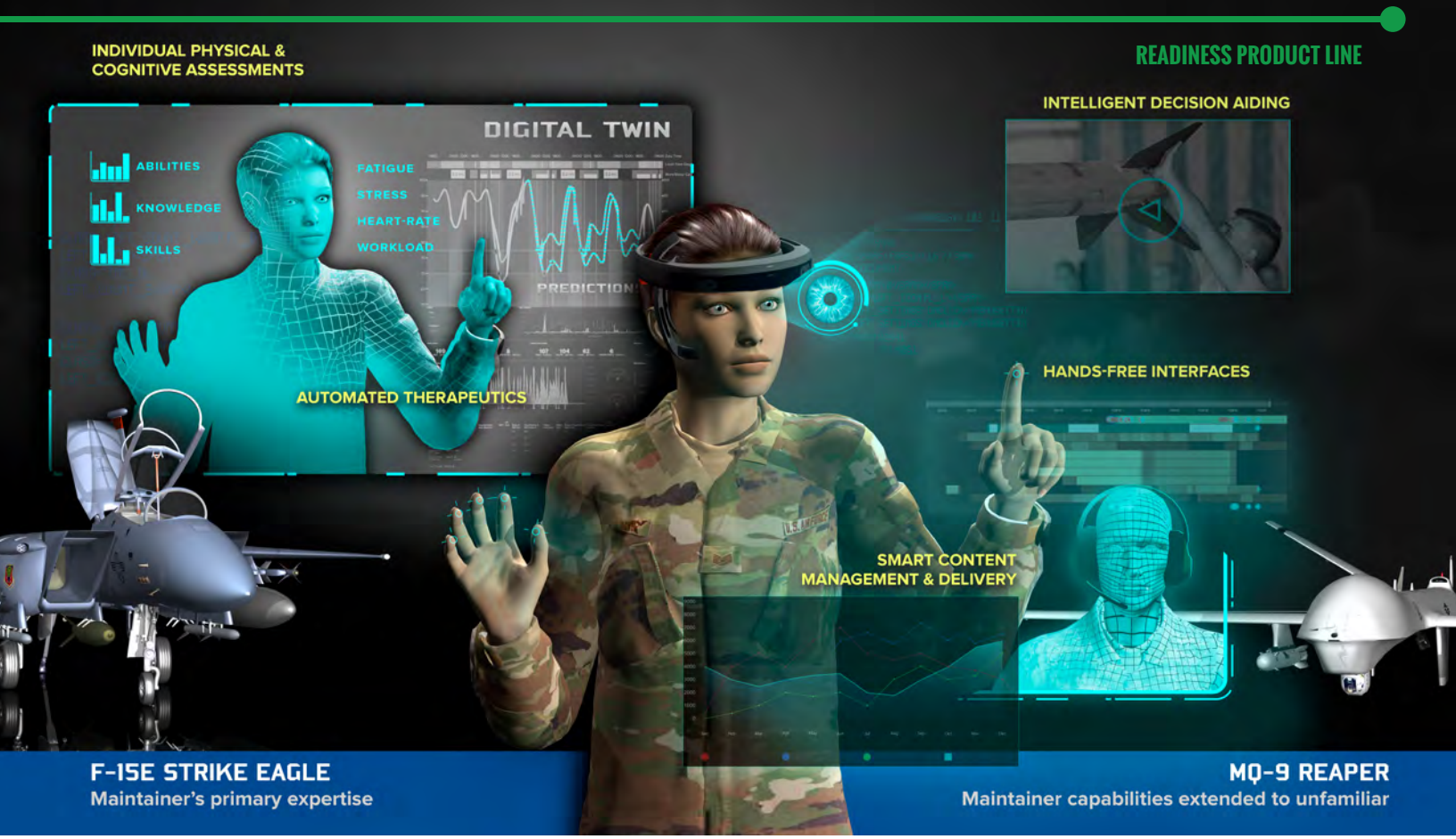
Integrate other CGF software into the Testbed: ASCOT-7 and Advanced Framework for Simulation, Integration and Modeling (AFSIM)



TESTBED

The Testbed utilizes many government-off-the-shelf (GOTS) software, such as Next Generation Threat System (NGTS), Network Integrated Constructive Environment (NICE), Tactical Observation Agent (TOA), Performance Evaluation and Tracking System (PETS), and Live, Virtual, and Constructive (LVC) Network Control Suite (LNCS). Aptima developed the Model to DIS (M2D) Server and the M2DIS Clients to allow the cognitive models to seamlessly interact with the Testbed without deep knowledge of the Distributed Interoperability Simulation (DIS) protocol. To achieve this seamless interaction, numerous external Application Programming Interfaces (APIs) were developed. Testbed APIs include scenario control (queue, dequeue, and run time), entity data retrieval (location, orientation, radar tracking list, weapon load, and fuel status), and entity controls (maneuvers, weapon fire, and radar state). The Testbed is CGF agnostic, so that any CGF can be easily integrated and used. ☆

Dr. Winston "Wink" Bennett, Readiness Product Line Lead, 711 HPW/RHW
Dr. Samantha Perry, Product Manager, Aptima
Dr. Luke Waggenspack, Researcher, Aptima
Dr. Will Dupree, Research Engineer, Aptima
Mr. Jacob Smith, Testbed Engineer, Aptima



Graphic by 711 HPW/RHW

JUST IN TIME:

Updates for Just-in-Time Multimission Airmen/Warfighters (JITMMA/W)

As our Air Force conducts operations according to Agile Combat Employment (ACE) concepts, in environments with austere forward operating sites, deploying large numbers of personnel with all needed skill sets will not be feasible. Instead, these environments will be characterized by reduced manpower, a smaller logistics footprint, and unreliable data connections. Our multi-capable Airmen (MCA) will have to perform novel and infrequent tasks under pressure and time constraints, often outside the usual limits of their specialties or career fields. MCA must have just-in-time, point-of-need support for training, performance, and resilience. To overcome this critical challenge to effective combat power generation in ACE and to support our multi-capable Airmen, the Airman Systems Directorate's Warfighter Interactions and Readiness Division (711 HPW/RHW) recently launched a project to create a training content creation, delivery, and assessment infrastructure. It will leverage Commercial-off-the-Shelf (COTS) and Government-off-the-Shelf (GOTS) technologies wherever possible and will deliver evolving capability through a series of technology evaluations and field experiments. While there are many possible use cases for this technology, the initial focus is on maintenance in austere locations. RHW is using model-based systems engineering processes and practices to design an agile and resilient infrastructure to integrate best-of-breed technologies in aiding/training agents, multimodal interfaces like speech recognition, and mixed reality (MR) technologies to deliver the right training content to the right Airmen at the right time. In addition, we plan to integrate unobtrusive physiological sensing and analytics to assess training, task performance, and Airman state to maximize results.

Our team of stakeholders, including representatives from AFRL, Air Education and Training Command (AETC), Air Combat Command (ACC) Agile Battle Lab, and industry partners, have observed ACE exercises and provided exemplar technology demonstrations to maintenance Airmen, leadership, and subject matter expert (SME) observers to garner insights about the usability and utility of these technologies in a deployed, austere, ACE environment. Regular field evaluations provide Airmen and Guardians with awareness of our efforts and enable our team to develop critical technology well-aligned to the users' needs. Field evaluations and technology assessments are fundamental components of our approach as we rapidly and iteratively create this critical enabling technology for our multi-capable Airmen and Guardians. ☆

Mr. David Malek, Senior Research Psychologist, 711 HPW/RHWL

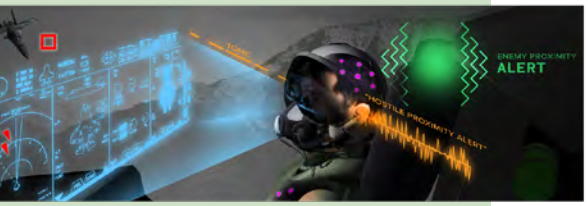
Dr. Deirdre Mahle
Airman-Machine Integration Lead,
711 HPW/RHW

AIRMAN-MACHINE INTEGRATION (AMI) PRODUCT LINE

Welcome back to the fall issue of Fight's On! The Airman-Machine Product Line (AMI PL) team has spent the past year focusing on strategic drivers. The demand signals have been nothing if not crystal clear about the needs of the AF. We have been tracking Air Force Futures' Top Ten for FY23, specifically Agile Combat Employment, Autonomous Collaborative Platforms, Base Defense and High Speed VTOL. The USDR&E Critical Technology areas encompass our core capabilities of Trusted AI and Autonomy, as well as Human Machine Interfaces. Perhaps the most important demand signals, though, are the SECAF's Operational Imperatives. We have been heavily engaged in:

- **OI#2** – Operationally Optimized Advanced Battle Management System (ABMS)/ AF Joint All-Domain Command and Control (JADC2)
- **OI#3** – Next Generation Air Dominance Family of Systems
- **OI#5** – Resilient Basing, Sustainment, and Communications in a Contested Environment

From our perspective, the overarching demand from all three is the need for modernized Battle Management Command and Control (BMC2). The current models are simply too vulnerable for a peer adversary and too slow for the future battlefield. With our adversaries having equally capable technology, it will be our decision superiority that will be the decisive factor of any conflict. In response to these demand signals, we have pivoted to JADC2. Our research teams have developed several new efforts to increase our capabilities dramatically to provide resilient and agile decision making in a distributed, yet digitally connected all-domain environment. Human-machine teaming, where we seamlessly synergize human cognitive capability with intelligent machine teammates, will provide the speed needed to turn vast amounts of data into decision advantage. The AMI PL will add three new efforts to the roadmap:



Graphics by 711 HPW/RHW

JADPACT (Joint All-Domain Planner with Adaptive Collaborative/Control Technologies) – a multi-domain command station with AI interaction/ cooperative planning and control (leveraging IMPACT - Intelligent Multi-UxV Planner w/Adaptive Collaborative Control Technologies), enabling JAD distributed operations. JADPACT will be include an ABMS compliant testbed and architecture to speed tool transition.

Next Gen Mission Planning & Debrief – mission planning HMI augmented with AI/ML algorithms to enable distributed, real-time planning capabilities, as well as map drawing tools that will transition to Web-based Information Dominant Warfare (WIDOW) (official AF Mission Planning Cell Tool).

Legion – a portfolio of efforts providing cloud based, Secure Internet Protocol Router/ Joint Worldwide Intelligence Communications System (SIPR/ JWICS) hosted capability to optimize the ISR process.

Our teams continue to connect across AFRL on several transformational, game-changing technologies, including Mayhem (multi-mission hypersonic platform), High Speed Vertical Take Off and Landing (a key capability for runway independence and ACE), and the 'wildly successful' Skyborg which will expand into Collaborative Combat Aircraft (CCA). I'm looking forward to another productive year as we rise to meet the challenges of the future. ☆

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

ANALYTIC TOOLS PRODUCT LINE

PIVOTING TO REMAIN RELEVANT

In response to AFRL's direction to align with the Secretary of the Air Force (SECAF) Operational Imperatives and to expand support to include the US Space Force, RH has made local pivots that: 1) implement Digital Engineering as a tool to model and predict biological and cognitive human performance, and 2) build products that empower US Decision Superiority against Near Peer warfighters and technologies. These pivots in RH will enable Airmen and Guardians to sustain higher levels of performance in all-domain, joint warfare environments that are saturated with information and technology.

The Analytic Tools (AT) Product Line is currently assessing its ability to provide analysis capabilities that leverage Digital Engineering in the lab to achieve Decision Superiority in the battlespace. From an analytics standpoint, Digital Engineering provides the ability to model the workflow, decisions, behaviors, skills, and attributes that provide unique insight into current and future performance of military personnel. Just as Finite Element Analysis or Computational Fluid Dynamics are industry-standard digital tools to understand and predict the response of natural and man-made structures to external stimuli, RH intends to grow and refine its implementation of Digital Engineering in an analogous way with humans. In fact, RH has been implementing a form of Digital Engineering for years through the efforts of Dr. Wink Bennet and Dr. Leah Rowe. By building operationally representative environments that enable the real-time interaction of Live, Virtual, and Constructive elements (e.g., humans, buildings, platforms, payloads, terrain, etc.), they have created the infrastructure needed to model humans and their interactions with technology.

To provide a clear slate for generating new products, AT divested of legacy efforts in FY23 to provide resources for newer priorities. Some of the existing tools, such as Haystack (a multi-media, multi-language translation tool) are being re-envisioned to expand focus from post-event Language Translation to real-time Verbal, Written, and Non-Verbal Communication Analysis. This potential pivot is based on the recognition that the ability to deeply understand human communication is a critical step to realizing the full promise of autonomy.

AFRL and the Technical Directorates have set the rudder and the internal race to relevance is on. Our ability to self-assess and pivot to remain relevant are critical in the current crossroads where global stakes are perceived to be high. ☆

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

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

> **Our ability to self-assess and pivot to remain relevant are critical in the current crossroads where global stakes are perceived to be high."**
>
> **— Mr. Greg Beister**
> Analytic Tools Lead, 711 HPW/RHW

TRAINING
CORE TECHNICAL COMPETENCY (CTC)

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

"Every Guardian must develop joint warfighting mastery to excel in complex and uncertain operating environments. A team-centric culture will enhance our ability to integrate seamlessly into joint, interagency, and coalition efforts... Our model for talent development and employment moves beyond generalized development to an individualized approach that provides every Guardian what they need, when they need it."

— U.S. Space Force, The Guardian Ideal

ADAPTING TO A FUTURE OF EVOLVING THREATS

As we arrive at the 2022 edition of Fight's On! for the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), change continues to influence every aspect of the Training Core Technical Competency (CTC) research portfolio, strategy, and culture. For many years we have emphasized adaptivity, flexibility, and agility in our vision for the future of Air Force readiness. However, that vision now more fully recognizes that these attributes are not capabilities we need to attain, but rather a long-term posture we need to adopt to ensure that our research continues to be responsive to both current requirements and longer-term operational imperatives.

As we evolve to meet these demands, our vision has expanded to allow our research to meet a broader set of demands for capabilities that ensure the readiness, problem solving capacity, and agility of our Airmen and Guardians to meet the needs of future operations. The focus is to allow the Department of the Air Force to fully leverage the cognitive strength and resolve of our people to maintain our edge in operations. This is reflected in our vision:

> Enable more lethal Air and Space Forces by creating, demonstrating, and delivering cognitive technologies for digital engineering and agile readiness that ensure mission effectiveness.

Continued on next page

Graphic by Mr. Will Graver

This vision supports the strategic orientation of the Air Force and the technology needs that will sustain our unfair advantage even as we face increasingly capable peer adversaries. We break this vision down into 3 goals, which reflect our broadened perspective on how to deliver capabilities to the Air Force that enable and enhance mission effectiveness:

1. Enable personalized, proficiency-based training to support multi-capable Airmen and Guardians in joint all-domain operations

2. Maximize mission effectiveness by establishing a persistent, global readiness ecosystem for individuals and teams, including teams of humans and machines

3. Advance considerations of human performance in system development and operational planning with cognitive & behavioral models for digital engineering

These goals emphasize our role in supporting future training infrastructure to realize the potential of shifting toward quantitatively-based metrics, models, and algorithms that allow for real-time assessment and prediction of individual readiness. We continue to recognize the critical importance of providing training where and when it is needed, with reduced manpower costs and security that provides opportunity for training at the joint and coalition levels on demand.

Finally, our broader focus envisions how our core research areas (CRAs) can contribute beyond the training arena; working collaboratively across the Airman Systems Directorate to enhance digital engineering environments that better represent the capacities and limitations of human performance in the range of predicted future operational environments, domains, and scenarios.

Embedded in our technical approach are emphases on four technology areas.

1. **Technology**: Innovative technologies are emerging in various sectors of industry at a rapid pace. We have a responsibility to identify what technologies offer opportunities to improve the pace, magnitude, and/or durability of training benefits and improvements in mission effectiveness.

2. **Data**: To meet long-term strategic objectives for readiness and multi-capable Airmen and Guardians, we must advance our capabilities to support persistent, high-resolution human and system measurement and quantitative, proficiency-centric readiness assessment and prediction at the individual and team levels.

3. **Team Training**: Future operations will be fundamentally integrated through joint and coalition all-domain operations. Individual assessments of readiness will remain critical; however, we must also develop and demonstrate joint learning environments and performance metrics that facilitate training, mission preparation, and rehearsal for teams of humans, with and without machine teammates.

4. **Agility**: As we prepare for a future battlespace that will be increasingly volatile, uncertain, complex, and ambiguous, we must create an ecosystem that prepares our Airmen and Guardians for rapid integration of new technology, provides for adaptation of training content to real-time operational data, and establishes the foundation for adaptivity, creativity, and innovation to meet the demands of tomorrow's operations.

The Training CTC has long led the Air Force in research and application to define the art of the possible in maximizing readiness. As we continue to advance, the Training CTC portfolio will also evolve – that is, change – to meet both short- and long-term priorities. To advance our science and technology, we will also continue to maintain and grow our enterprise-wide communication, leveraging, and collaboration to meet the complex needs for the future force. ☆

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

COGNITIVE MODELS (CM)
CORE RESEARCH AREA

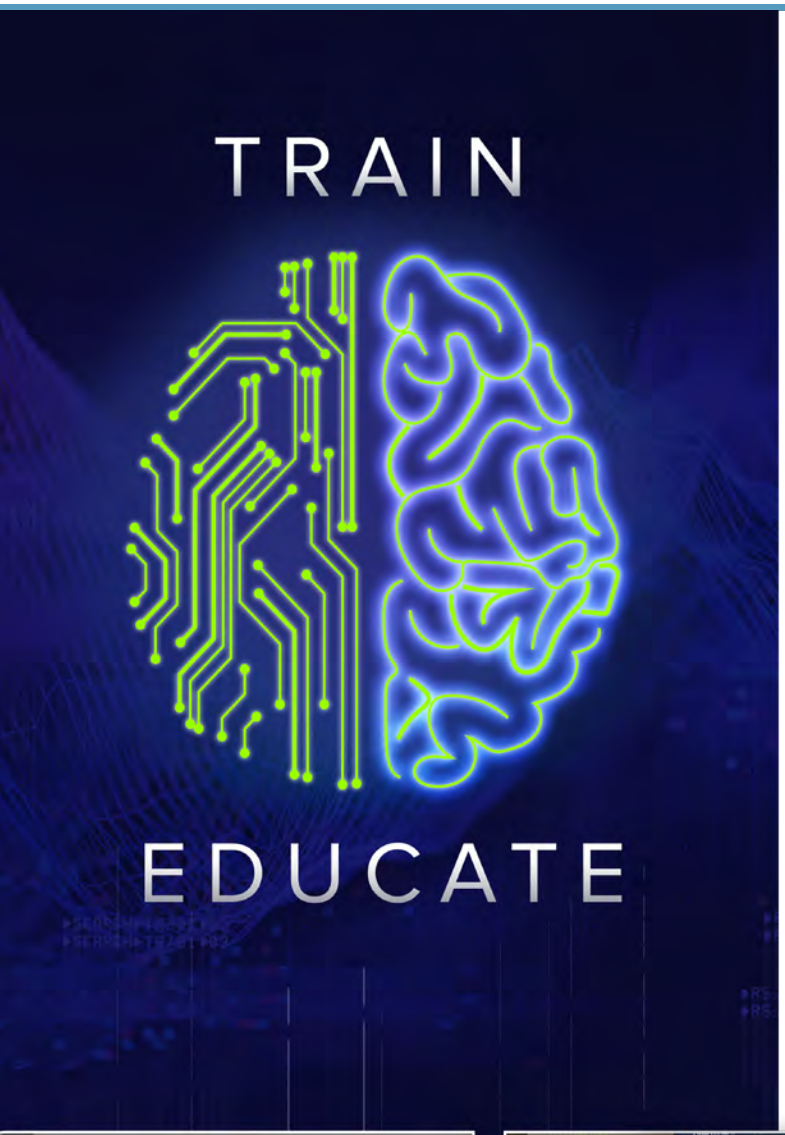
Dr. Megan Morris
Cognitive Models Core Research Area
Lead, 711 HPW/RHWM

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) 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 a diversity of 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, 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 with the goal of providing real-time, behavior-specific assessments and predictions of readiness based on stressors such as fatigue, sustained attention, workload, pharmaceuticals, toxins, and environmental factors. This past year our team has emphasized integrative capabilities within the lines of effort that focus on delivering technical capabilities that can be transitioned through product lines. The following articles highlight these endeavors. ☆

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



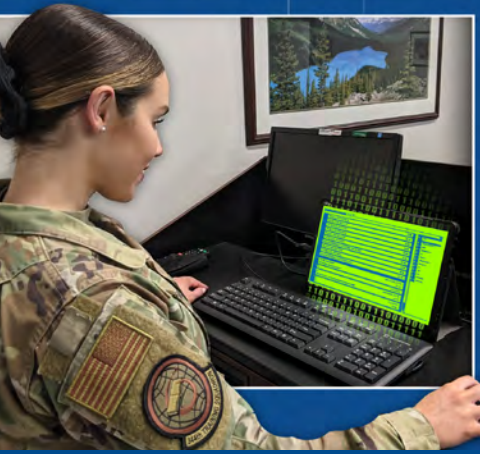
TEACHABLE MODELS FOR TRAINING

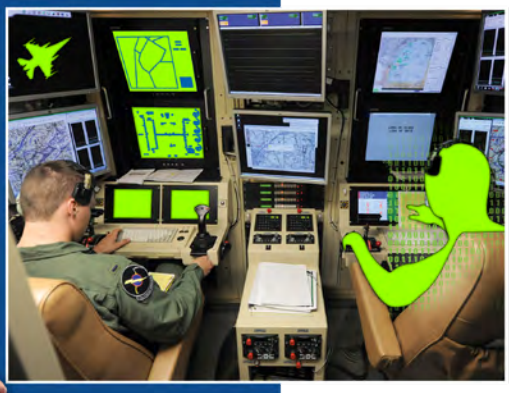
Developing intelligent systems capable of operating as team members, adversaries, or tutors within team training scenarios is costly and time consuming. A consistent issue is the requirement for developers to analyze, organize, and integrate task knowledge within a modeling-and-simulation framework (a.k.a. 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:

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

Graphic by Mr. Will Graver

Graphics by 711 HPW/RHW

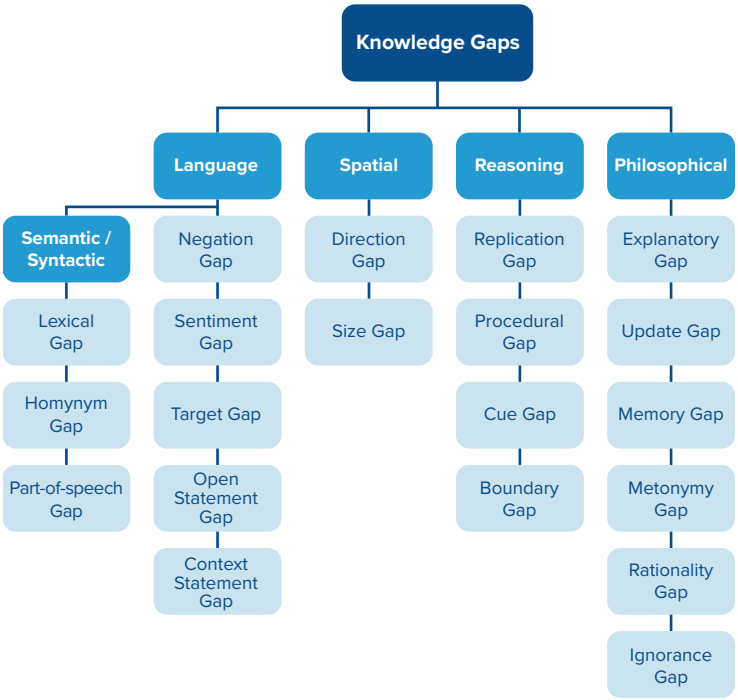
KNOWLEDGE GAP DETECTION, IDENTIFICATION, AND RESOLUTION

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 more important. Detection, identification, and resolution of gaps in knowledge required to successfully execute newly instructed missions will help to reduce autonomy failures and mistakes. Our research and development approach is that it is preferable to identify inconsistencies and absences within a knowledge base early, rather than assuming 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 even 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 is 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 & resolution (Bajaj, et al, 2020; 2021) have been investigated. ☆

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

ADAPTIVE LINGUISTICS

For decades, research on human-computer interaction has occurred where the focus was 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 DoD goals. The ability to seamlessly team with intelligent machines requires advancements in machines' capabilities to adapt to changes in communications. The Adaptive Linguistics task is focused on research investigating how to expand the language capabilities of current cognitive models and synthetic teammates to include the ability to incorporate 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 use to each other to accomplish a particular task, and will additionally leverage 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 which 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.

Over the past year, development of a cognitive model with adaptive language capabilities has been ongoing. We have also been preparing to evaluate the impact of such a model on the performance of human-machine teams engaged in a collaborative card task. We submitted and were successfully awarded an Air Force Office of Scientific Research (AFOSR) Laboratory Research Initiation Request (LRIR) Lab Task: The Role of Structural Language Adaptation in Human-Machine Team Performance. The Lab Task will fund a series of experiments, the data from which will be used to directly test the hypothesis that a cognitive model with adaptive language capabilities is not only preferred by human teammates but also contributes to an improvement in human-machine team performance when compared to a model without such adaptive language capabilities. ☆

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

Graphic by 711 HPW/RHW

RAPID MULTI-SKILL ACQUISITION IN MODELS

The Air Force Research Laboratory (AFRL) is working with the Air Force Office for Integrated Resilience (AF/A1Z) to fundamentally transform annual training requirements for Sexual Assault Prevention and Response (SAPR) and bystander intervention training. This collaborative total force overhaul will shift away from the "one-size-fits-all" classroom-based modality that has historically tracked attendance rather than competencies, and move towards a proficiency-based assessment and tailored learning approach. These assessments will require all Air Force personnel to apply their knowledge to scenario-based training videos embedded within the My Learning platform starting in January 2024. This paradigm shift will execute a phased approach.

The first phase will shift to proficiency-based performance assessments and tracking to generate a wealth of data and better understand the interindividual variability of learning and forgetting dynamics within this training domain. The second phase will personalize, optimize, and tailor learning content to individuals on the basis of unique learning needs, using vetted, validated, and patented cognitive and artificially intelligent (AI) predictive analytics developed by AFRL. In this way, we will develop a precedent and template for a depth and breadth of other Air Force total force training applications to shift towards the gold standard of proficiency-based, personalized learning. ☆

Dr. Tiffany Jastrzembski, Senior Cognitive Scientist, 711HPW/RHWS

PHASE ONE

Shift to proficiency-based performance assessments and tracking to generate a wealth of data and better understand the interindividual variability of learning and forgetting dynamics within this training domain.

PHASE TWO

Personalize, optimize, and tailor learning content to individuals on the basis of unique learning needs, using vetted, validated, and patented cognitive and artificially intelligent (AI) predictive analytics developed by AFRL.

Photo by Tech. Sgt. Daniel Martinez

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

A collaboration between Airman Systems (RH) and Aerospace Systems (RQ) has produced improvements to the cognitive fidelity of agents within the Advanced Framework for Simulation, Integration, and Modeling (AFSIM). Typically, AFSIM 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; the agents do not accurately reflect human cognitive capacities.

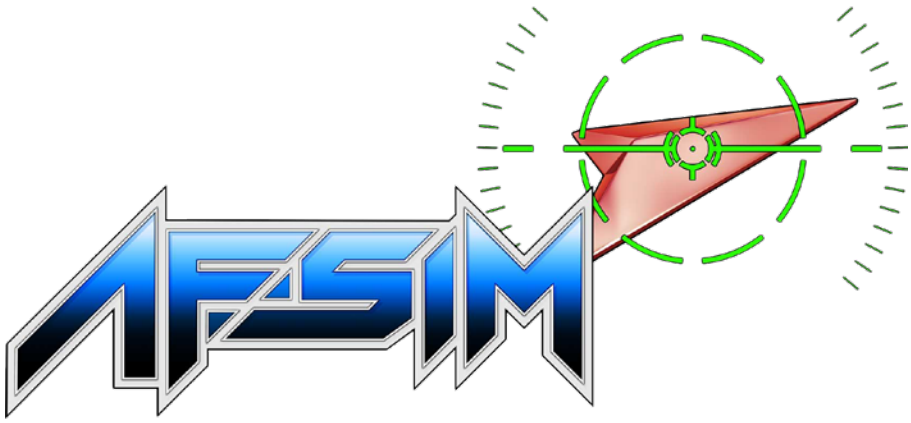
The RH-RQ team has focused research and development on limiting agent capabilities within the Observe and Orient components of the Observe-Orient-Decide-Act (OODA) loop; specifically, limiting the ability for agents to rapidly and completely process information for all tracks available to any particular agent to capabilities observed in humans. To this end, delays associated with perception, assessment, decision-making, and motor behavior have been incorporated into AFSIM. Directed-energy effects on cognitive performance, specifically, the effects of laser dazzle and flash blindness on perceptual encoding abilities while piloting aircraft, have also been incorporated into AFSIM. These new capabilities, part of AFSIM Release 2.9/2.7.4, are currently available to the end user. ☆

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

The team continues to improve AFSIM Agent cognitive fidelity through:

- **Evaluation of perceived risk**
- **Cognitive effects of fatigue**
- **Capacity limits on attention**
- **Capacity limits on maintained information**

Photo by U.S. Airman Shawna Keyes



MULTISCALE MODELS FOR COGNITIVE PERFORMANCE

Warfighters must execute challenging missions all over the globe. These operations put stress on their bodies and their minds. In the Multiscale Models for Cognitive Performance line of effort, we emphasize powerful, theory-driven technologies that integrate physiological, cognitive, and behavioral models to anticipate, detect, and mitigate effects of operational stressors on warfighter performance. This line of effort is developing promising technologies such as the Mission Readiness App, real-time assessment and modeling of operator stressors, and human modeling tools for wargaming.

The Mission Readiness app is an inclusive tool allowing warfighters, supervisors, and mission planners to reduce risks related to operational stressors. It is a smartphone application in which a warfighter's sleep history and caffeine consumption can be tracked. Based on this data, the app projects a readiness profile throughout the following day, allowing for high-impact tasks to be planned during optimum readiness windows and periods of low readiness to be avoided or mitigated. Moreover, these projections of readiness are updated in real time based on new data, ensuring that the most accurate projections are possible under current mission conditions. These projections are refined even further by a real-time behavioral metric known as the Psychomotor Vigilance Test, a simple reaction time test that can be administered from within the app.

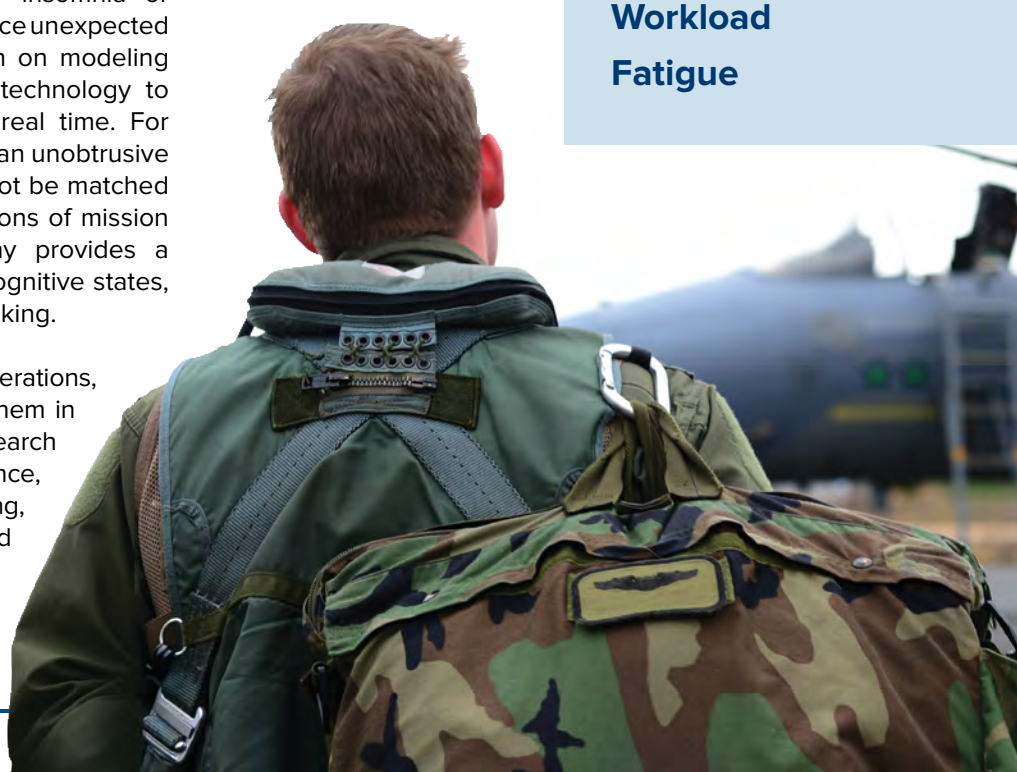
A significant challenge associated with minimizing operational risk is accounting for rapidly changing operational conditions. Crew rest schedules can be disrupted by insomnia or unscheduled interruptions, or a task may introduce unexpected cognitive stress. Therefore, we lead research on modeling of operational stressors and use wearable technology to update the projections of those models in real time. For example, wearable actigraph watches provide an unobtrusive measure of sleep length and quality that cannot be matched by self-report, enabling more realistic projections of mission readiness. Moreover, electroencephalography provides a powerful tool for detecting mission-relevant cognitive states, including fatigue, workload, effort, and multitasking.

Warfighters engage in a wide variety of operations, so our technology must be ready to support them in any environment. For this reason, we lead research in cross-domain readiness and performance, developing models that generalize across piloting, vigilance, and intelligence, surveillance and reconnaissance tasks. In an ongoing effort, we are using a combination of wearables, cognitive modeling, task analysis and modeling, and cognitive task batteries to predict fatigue and cognitive load in operators in C-17 operations and a naval watch task. In this work we aim to create a set of capabilities to describe, assess, and predict the cognitive and physiological demands experienced by operators across environments and operations.

Wargaming is a powerful tool for training and operational planning, but operational demands on humans are usually represented at a very low level of fidelity, if at all. This results in potentially unrealistic scenarios, and leaves wargaming users unable to simulate the effects of fatigue and other operational stressors on crews and operators. Therefore, we have been developing models that can make meaningful predictions of warfighter performance in operational engagements. These include models of fatigue in maintenance crews based on rest and work schedules and models of accuracy and response time in radar operators as a function of task fatigue and electronic warfare. ☆

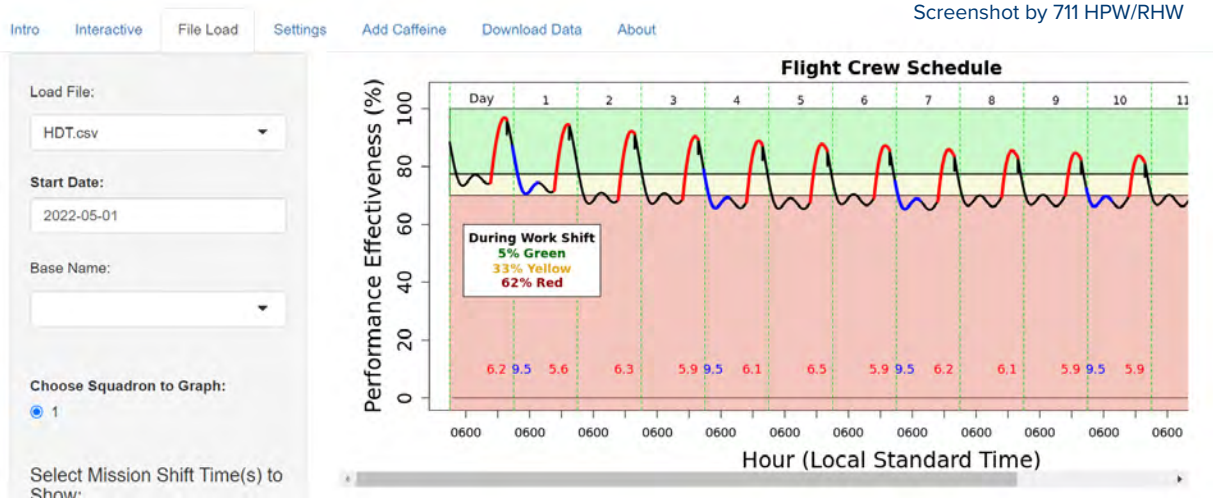
Dr. Christopher Stevens, Research Psychologist, 711 HPW/RHWM

Pharmaceuticals
Hypoxia
Toxins
Workload
Fatigue



Graphic by Mr. Will Graver

Screenshot by 711 HPW/RHW



FATIGUE AND SUSTAINED ATTENTION PERFORMANCE IMPACTS

Caption: A screenshot from the Fatigue Modeling in Wargaming Application. The image shows fatigue estimates in the form of a performance effectiveness line throughout several days based on information from an iSWAT Air Tasking Order. Shown is a realistic mock-up of a mission scenario for a squadron.

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 Air Force. Specifically, we focus on developing models that produce individualized, real-time, behavior-specific estimates of cognitive performance to inform mission readiness.

Our team has had great success this past year and we are increasingly working toward more integrative efforts with the other tasks within the Multiscale Models for Cognitive Performance Line of Effort to develop more holistic models of fatigue effects on performance.

We had several conference papers this year from our basic research portfolio. We presented two papers at the virtual and in-person International Conference on Cognitive Modelling (ICCM) 2022. The first focused on the development of a cognitive model predicting performance for simultaneous and successive discrimination vigilance tasks. The second focused on the development of a cognitive model that integrates electroencephalogram data, specifically gamma power, to increase predictive accuracy of performance in the Psychomotor Vigilance Test (PVT). We also presented at the 2022 Human Factors and Ergonomics Society (HFES) International Annual Meeting. The effort used Granger Prediction to evaluate fatigue effects on network gamma connectivity during the PVT. We also published a paper in the International Journal of Human Factors and Ergonomics focused on teasing apart the cognitive mechanisms underlying the vigilance decrement during simultaneous and successive discrimination tasks.

We published an article in Aerospace Medicine and Human Performance focused on comparing fatigue estimates from prescriptive sleep schedules generated in a fatigue risk management tool to fatigue estimates from actual sleep (via actigraph) in C-17 aircrew during operational missions. We also presented work at the Aerospace Medical Association (AsMA) 2022 Annual Scientific Meeting focused on comparing sleep and fatigue estimates between a commercial fitness watch and research-grade actigraph watch.

Our team continues to enhance our mobile fatigue application, the Mission Readiness App, with new capabilities. The app is now integrated with the Garmin Connect™ Application Programming Interface (API) and we are currently developing the capability to integrate with the Garmin Companion Software Development Kit (SDK) to enable even richer physiological data collection from Garmin devices to calculate sleep estimates. We recently presented work at AsMA 2022 discussing further enhancements to an individualized implementation of a biomathematical fatigue model and the integration of this individualized model into the Mission Readiness App. We are also integrating caffeine modeling and the capability to provide fatigue mitigation countermeasure suggestions based on this modeling into the app through collaborative efforts with the Physiocognitive Modeling for Environmental Stressors team.

We will be presenting work at the 2022 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) detailing our development of an application that models fatigue based on aircrew and maintenance personnel information from the Integrated Sustainment Wargaming Analysis Toolkit (iSWAT), a wargaming logistics simulation software. We are currently working on incorporating caffeine effects into the model based on work by our collaborators, Johns Hopkins University's Applied Physics Laboratory. ☆

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

INDIVIDUALIZED COGNITIVE LOAD PROFILING

Warfighters face demanding, complex, and ever-changing operational tasks. The Individualized Cognitive Load Profiling task works to predict operational performance of warfighters by creating models of the cognitive demands they experience during their missions. We are currently examining four domains: Intelligence, Surveillance, and Reconnaissance (ISR; Fisher, Frame, & Stevens 2022), 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; Swan, Stevens, Fisher, & Klosterman, 2022).

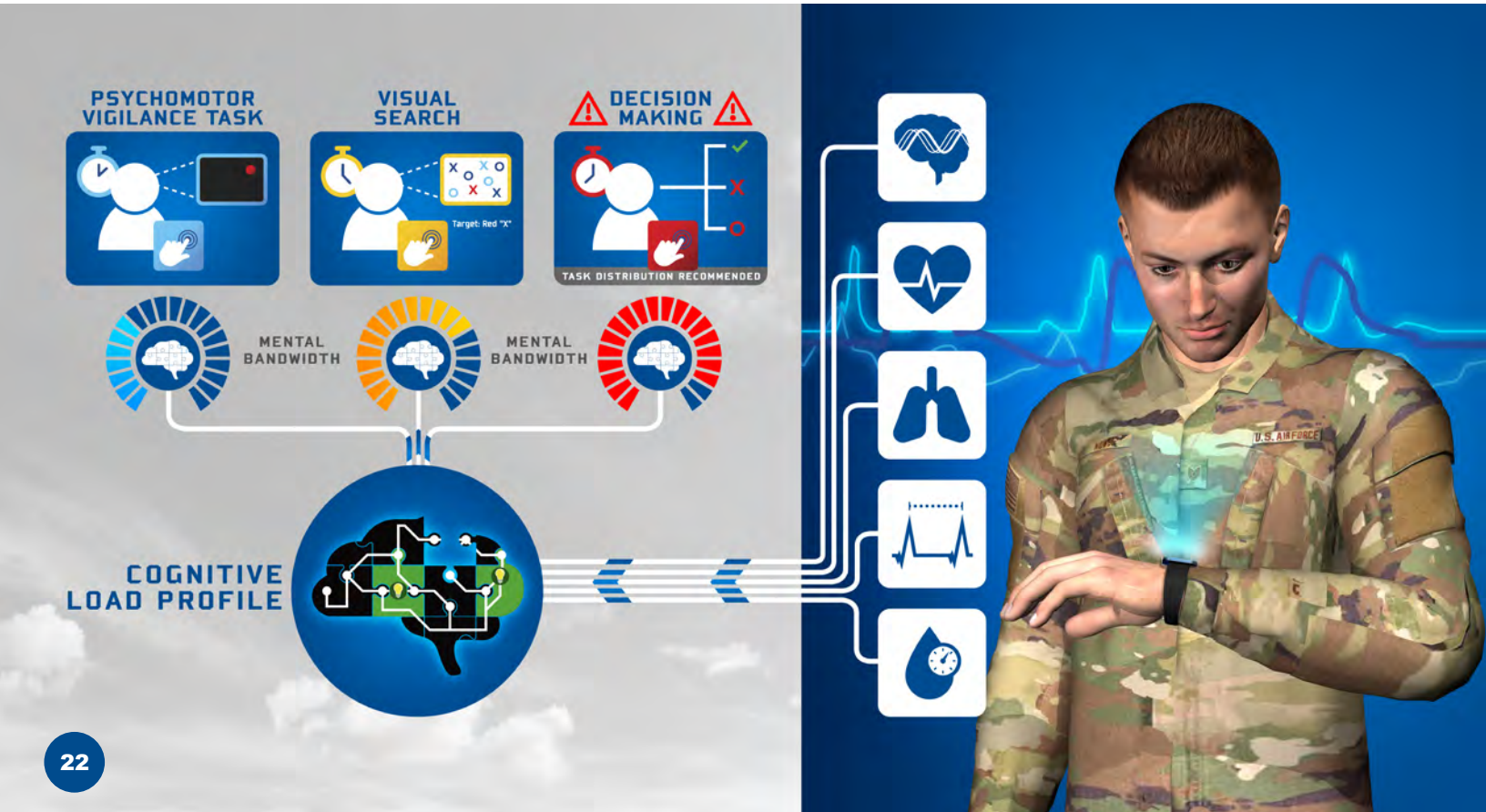
Previously, we have demonstrated agreement between our cognitive models and estimates of global workload obtained from physiological metrics (Stevens, Morris, Fisher, Myers, 2019), and subjective reports (Stevens, et al., 2019; Stevens et al., 2020). In these tasks, we are also examining the effects of individual differences, such as working memory capacity and personality traits, on workload. These models will provide important information about warfighter tasking decisions, such as whether additional tasks can be added without sacrificing performance. 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 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 simulation exercises, 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., 2022). ☆

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

The task is currently examining four domains:

1. **Intelligence, Surveillance, and Reconnaissance**
2. **Aerial Refueling**
3. **Command and Control**
4. **Multitasking in Aircraft-related Tasks**

Graphic by 711 HPW/RHW



PHYSIOCOGNITIVE MODELING FOR ENVIRONMENTAL STRESSORS

Warfighters' cognitive performance on tasks can be hindered through the 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 *phsysiocognitive modeling* effort 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, 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 ensuring 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 focus of much of the empirical and modeling research (see Integration function, Figure 1). 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 caffeine effects interacting with fatigue (see Figure 2 ; Halverson, et al., 2021; 2022). 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 helping to determine if, and to what extent, cognitive capacities are affected differently by exposure to different compounds. We are currently investigating approaches to modeling and predicting the effects of hypoxia on cognitive performance with partners from the University of Florida and the Mayo Clinic, as well as collaborating with the 711th Human Performance Wing's Airman Biosciences Division to collect data from Airmen exposed to isopropanol in isolation, and plan to collect data on G-forces effects on cognitive capacities. ☆

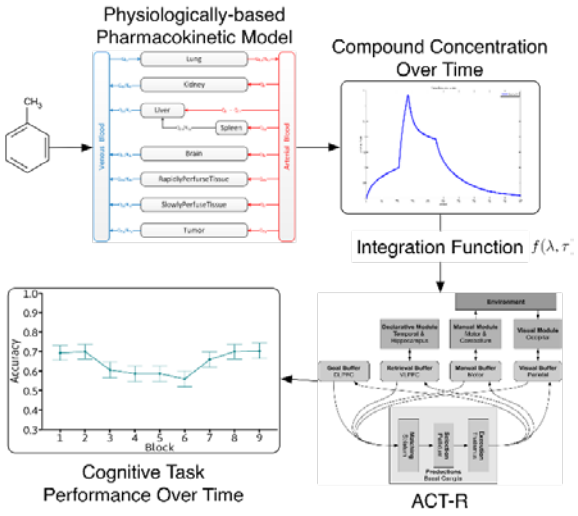


Figure 1. Physiocognitive modeling process.

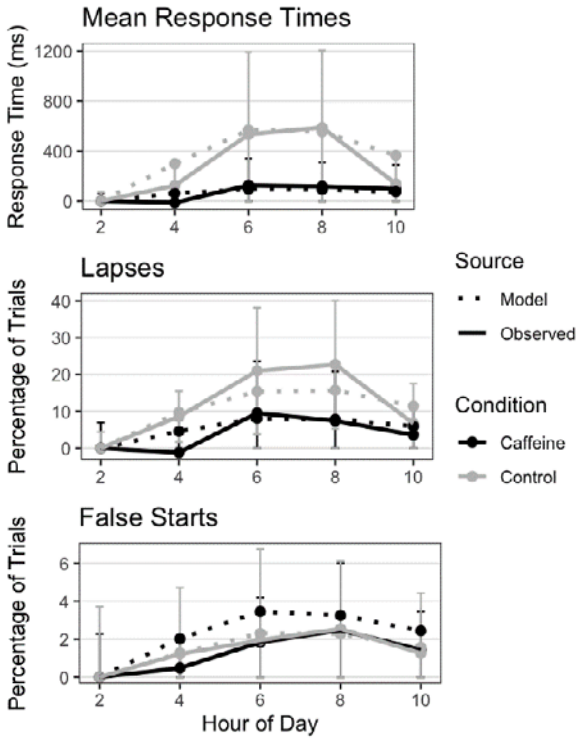


Figure 2. Caffeine physiocognitive model performance against human data in terms of response times, lapses, and false starts across caffeine and control conditions.

Graphics by 711 HPW/RHW

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

THE GAMING RESEARCH INTEGRATION FOR LEARNING LABORATORY®

Photos by Mr. Will Graver

FULL THROTTLE STEM

The Full Throttle STEM (FTS) event was back in full swing this year with two locations for twice the excitement. Students from the surrounding areas of Darke County, OH made their way to the Eldora Speedway for the annual FTS hosted by the GRILL®. Hundreds of students moved from one station to the next viewing different science exhibits, gaining insight into other schools' science programs. The newest station on display this year was drone racing. With industry and different branches of the military relying more on Unmanned Aerial Systems (UAS), this was a great chance for a hands-on experience piloting aircraft around twists and turns and even through hoops. The event culminated with the main attraction; students from different schools spent the year designing remote controlled cars with varying weight distribution, axle length, and plenty of other variables to see which design would stand out as the fastest on the track. In years past, FTS would end here, but this year the event was also held at the National Museum of the United States Air Force. The event was originally designed to drive interest in science, technology, engineering, and math (STEM) for kids in rural areas, but this year Dayton students got to view the exhibits and develop their own racecars. The GRILL® was at both sites to put their own simulators on display for the kids to play with and ask questions about becoming software developers. Both events were a resounding success with kids and adults alike getting to enjoy STEM in action. ☆

Capt. Eric Lawson, GRILL Program Manager, RHWM

VISION

The Gaming Research Integration for Learning Laboratory (GRILL®) leverages low-cost, high-fidelity, commercial off-the-shelf technology to create a variety of solutions for partners in the Department of Defense (DoD).

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.

"(Full Throttle STEM helps) to ensure the development of a STEM-oriented workforce that can embrace and overcome the advanced challenges of the future."

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

GRILL® DIS PLUGIN

Interest in Unreal Engine for modeling and simulation has accelerated recently with the announcement of Project Anywhere and 64-bit floating point precision support in Unreal Engine 5. Another critical link in the chain for the modeling and simulation community is support from within Unreal Engine for the Distributed Interactive Simulation (DIS) protocol. DIS has long been an Institute of Electrical and Electronics Engineers (IEEE) standard embraced throughout the military defense enterprise. DIS provides a protocol by which simulations can interact in an ad-hoc and "plug and play" fashion.

The GRILL® has developed this needed DIS capability for Unreal Engine with their GRILL® DIS for Unreal plugin. To bring DIS into Unreal Engine, the GRILL® utilized Open-DIS, the free, open-source implementation of DIS maintained by the military's Naval Postgraduate School. Initially developed for Iron Dev at I/ITSEC 2021, the GRILL® has expanded the plugin and demoed it publicly for the first time at the Geospatial Intelligence (GEOINT) Symposium in support of the Cesium booth.

This DIS for Unreal plugin provides basic User Datagram Protocol (UDP) socket support and handles interpreting of DIS network traffic using the Open-DIS libraries. Then, the plugin uses the interpreted DIS data to create and update user specified DIS entities in Unreal Engine. The DIS plugin provides an interface for a user to add on custom functionality when specific DIS data is received. The plugin also supports sending DIS information through utilizing the Open-DIS libraries allowing DIS data to be sent out of Unreal Engine to other DIS-capable simulations.

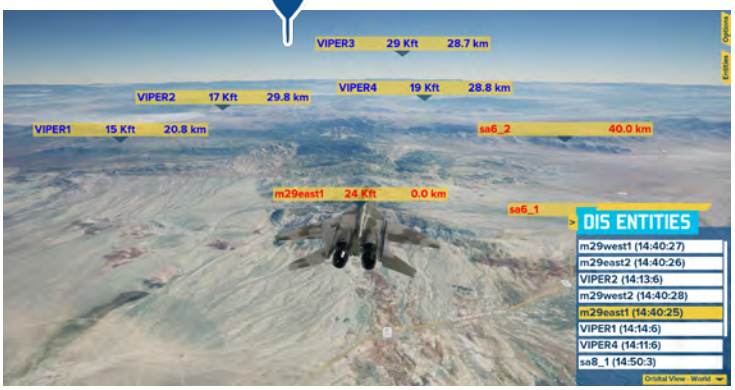
Core components of the GRILL® DIS for Unreal plugin are decoupled to allow for custom functionality to be added easily. If custom UDP socket support, Protocol Data Unit (PDU) processing, or managing of DIS entities are desired, the basic supplied components can be replaced by new components. Furthermore, the GRILL® DIS for Unreal plugin is not tied down to a specific round earth model. The plugin provides its own geospatial conversion functions which utilize the Georeferencing plugin made by Epic Games.

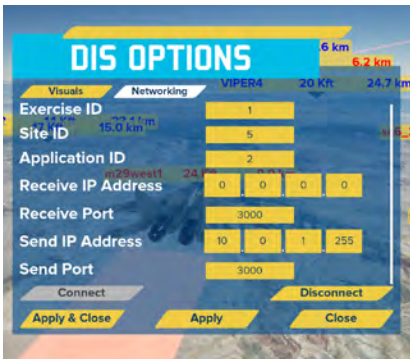
The GRILL® DIS for Unreal plugin currently has base functionality for a select few DIS PDUs. These include:

- **Entity State**
- **Entity State Update**
- **Remove Entity**
- **Fire**
- **Detonation**
- **Start Resume**
- **Stop Freeze**

The GRILL® has made their plugin open source on GitHub and invite the broader community to contribute to the plugin's development through implementation of additional PDUs.

Screen Captures by 711 HPW/RHW





The plugin can display external DIS entities being captured in Unreal Engine, as well as features such as pathing enabled to show dead reckoning capabilities. It also provides a variety of networking options.

The GRILL® is excited to announce that the GRILL® DIS for Unreal plugin is now available for download on the Unreal Engine marketplace for both Unreal Engine 4.27 and 5.0. A sample GRILL® DIS for Unreal project is also available for download from GitHub for you to explore and learn from. This sample project utilizes the Cesium for Unreal plugin for the global terrain and imagery. ☆

Mr. Jerry Huggins, Software Engineer, Ball Aerospace

Watch a Sample Video of the Plugin HERE

THE GAMING RESEARCH INTEGRATION FOR LEARNING LABORATORY®

SUMMER AT THE GRILL®

The GRILL® hosted its annual open house July 21st. High school students selected for the coveted Wright Scholar Program spent nine weeks learning the ins and outs of modeling and simulation. They developed a solution to an AF training need submitted earlier this year. The scholars learned how to communicate with their customer and amongst themselves to develop a product and meet goals. During the summer program they combined different technologies such as the gaming engine Unreal, Raspberry Pi, and virtual reality (VR) technology to accomplish their objectives. This year, 17 students and their college mentors developed five VR trainers that will serve as rapid prototypes for an AF training need. The prototypes developed included a live, virtual, and constructive (LVC) component for Calamityville, improvements to a laser dazzle effect simulator, a hands-fire extinguisher trainer, a simulator to assess reaction time and accuracy of laser eye protection, and a multiplayer roleplaying exercise that would allow participants to develop communication and team working skills. This program also allowed teachers to modernize their curriculum to better suit their students. The teachers demonstrated what they are doing to help students of today, be better prepared for the world of tomorrow.

A virtual rocket launcher or rocket-propelled grenade (RPG) within a simulated environment utilizing Distributed Interactive Simulation (DIS) network standard. The approach allows simulated data to be rendered in the virtual environment for testing and training.

A VR simulation that provides a realistic simulation of the effects to the visual system when a laser is shined at the launcher Dazzle and flare effects are simulated to provide training opportunities on their appearances and how Laser Eye Protection can dull these effects.

A multiplayer search and rescue simulator where members work together to rescue victims and address hazards in a collapsed building. The task requires collaboration among members and collects event-based data to research team effectiveness.

A simulated C-130J jet that simulates flight tasks and identifying aircraft symbols while simulating the effect of wearing laser eye protection (LEP) that can distort perception. The simulator provides the opportunity to research negative performance effects of LEP.

An immersive mixed reality game that allows maintainers to practice fire extinguisher procedures in a virtual environment (VE) while holding a real fire extinguisher that is tracked and rendered in the VE Data collected by the simulator can be used for training effectiveness research evaluations.

Photos by Mr. Will Graver | Screen Captures by 711 HPW/RHW

SUMMER AT THE GRILL® (CONTINUED)

US Air Force Academy (USAFA) cadets came to the GRILL this summer for an ongoing effort that will align with their capstone projects. The cadets are working to integrate a jammer/detection system built from a Raspberry Pi and commercial off-the-shelf (COTS) parts into an ergonomically controlled device. They will integrate the system with a VR application to develop the interactive control system as an intermediate step. The cadets are gaining valuable skills that enhance their educational experience while allowing faculty-cadet research. The project itself is just a small piece in a much larger mission of establishing an off-base facility near the academy to provide STEM outreach programs and provide VR/Augmented Reality (AR) solutions for Air Force training needs. This faculty-cadet research will enable current and future warfighters.

Lastly, local middle and high school teachers worked with the GRILL® team over the summer to develop STEM resources that are provided free on the lab's website. These included:

STEM-ply Summer Camp: Teachers hosted two camps for 6th-8th grade students who are passionate about STEM and getting hands-on learning. Students were introduced to the engineering design process through building a cardboard motor car. Information technology was introduced through video game design and Computer-Aided Design (CAD).

STEM Carnival: Carnival Project Schools hosts events throughout the school year to promote excitement for Science, Technology, Engineering , Art and Math (STEAM) topics for K-4 students. This carnival is a way to kickoff this effort by setting up multiple stations with demos and hands-on activities. These lessons will be made available to teachers to further promote STEM into their classes.

Flipped Classroom Guidance: Flipped classrooms are employed to enhance student comprehension of complex topics, by providing recorded lesson content for students to watch as needed. Teachers are able to complete practice problems in the classroom with teacher guidance. Students are then able to re-watch recordings anytime to meet their individual needs. ☆

Capt. Eric Lawson, GRILL Program Manager, RHWM

Photos by Mr. Will Graver

For additional information, visit the GRILL's® website:

WWW.AF-GRILL.COM

SUMMER POST-DOC & FACULTY FELLOWSHIP RESEARCH

DR. STEPHANIE FUSSELL
Kent State University

Researching the use of extended reality (XR) technologies for aviation and aerospace training, focusing on XR usability, user experience, and transfer of training. She is developing a methodology to compare VR head-mounted displays (HMDs) based on learning objectives, technical specifications, physiological impacts, and user experience.

DR. KELSEY MERLO
University of South Florida

Researching communication patterns in a cooperative dyadic virtual reality task. Dr. Merlo is working with GRILL® Engineers to develop a research-based simulator mimicking the popular VR-game Keep Talking and Nobody Explodes which sends data to common research software such as Noldus for expedited analysis.

MS. KENDALL CARMODY
Florida Institute of Technology

Researching dynamic human agent teaming (HAT) and how trust dynamics unfold in complex, heterogeneous, multi-agent HATs. Ms. Carmody is working with GRILL® engineers to design a testbed that supports human-agent interdependence and leverages elements from extant research to integrate innovative approaches.

PERSONALIZED LEARNING & READINESS SCIENCES (PLRS)

CORE RESEARCH AREA

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

The PLRS ("pillars") CRA is focused both on scientific discovery 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 individuals, teams, and teams of teams.

Graphic by 711 HPW/RHW

The technical challenges we are tackling within PLRS are organized around two themes: 1) Adaptive Proficiency Technologies (APT) and 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. Within Co-learning research, our top priorities include defining new paradigms for integrating emerging artificial intelligence (AI) technologies into intelligent tutoring, human-machine co-training and interactive learning, and establishing next-generation techniques for inference and mutual understanding between human and machine team members in a common mission environment. Within both lines of effort, digital training data are key enablers. The availability of digital data is critical for achieving high-resolution measurement and proficiency-centric readiness evaluation for individuals and teams, and enabling unit-level assessment 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 Research Psychologist, 711 HPW/RHWL

TWO PROJECT AREAS

- **Adaptive Proficiency Technologies**
- **Human and Machine Co-Learning**

ADAPTIVE PROFICIENCY TECHNOLOGIES

ERROR RECOGNITION AND RECOVERY TRAINING

The current DoD personnel and readiness strategy lays out the significant forces shaping the needs and challenges of our workforce. Among the goals of our strategy is a resilient and adaptive force capable of dealing with the evolving demands for skills and performance. From a training perspective, how do we support instructors to achieve this goal? How do we optimize curriculum and scenario content to maximize training effectiveness and readiness in an evolving future mission landscape? These questions motivated investment from the Personalized Learning and Readiness Sciences Core Research Area (PLRS CRA) in the area of error recognition and recovery (error management) training.

Our team began work in this area by conducting a targeted literature review, from which five key insights emerged:

1. There is a distinction between error avoidance training and error recovery training. The former focuses only on correct procedures, while the latter focuses on how to recognize and mitigate errors that may arise during mission performance.
2. Error recovery training has been found to be useful across a diverse range of domains and theoretical traditions.
3. Error recovery training may be best implemented in stages for some tasks, teaching learners how to learn correct procedures first, then adding error recognition and recovery.
4. Simulation-based exercises can be an important component of error recovery training, particularly when training for adaptive skill transfer.
5. Documented error recovery programs include a range of didactic, experiential, and reflection components.

From the literature of existing error recovery training programs, we extracted a set of design seeds for creating error recovery training. Next, our empirical work will begin to test the efficacy of these design seeds in the context of real-world training. We are gearing up to complete a study to test the error recovery training approach in the context of hemorrhage management training (tourniquet use). Our study goals include first establishing non-inferiority with a relevant target population (i.e., this training approach is not worse than status quo). For this evaluation phase, we plan to assess student recall of different errors and the ability to identify appropriate mitigating strategies. Next, we plan to evaluate whether the training helps prevent memory and skill decay. We hope to address skill decay through the unique focus on errors and our strong emphasis on using scenarios and narrative instruction.

Photo by Ms. Audrey Chappell

Research in this area addresses challenges associated with defining and quantifying readiness standards for individuals, and is a first step toward linking those with principled scenario content for learning and readiness validation. This work is unique to the portfolio of investment within the Training CTC due to emphasis on task context and complexity factors that make performance errors more likely to occur, as well as focus on scenario design principles to support the development of macrocognitive skills and expertise. ☆

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

PUBLICATIONS

Militello, L., Wagner, E., Winner, J., Sushereba, C., & McCool, J. (2021). Error recovery training literature review: Implications for emergency field medicine. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 65, No. 1). Sage CA: Los Angeles, CA: SAGE Publications.

Militello, L., Sushereba, C., Fernandez, R., Wolf, S., Winner, J. (2022). Learn, experience, reflect framework for training design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 66, No. 1). Sage CA: Los Angeles, CA: SAGE Publications.

Graphic by Dr. Julie Cantwell

OPTIMIZED LINGUIST CURRICULA AND VIRTUAL REALITY TRAINING LEARNING FOR DEFENSE LANGUAGE INSTITUTE (DLI) STUDENTS

Predictive Analytics for Learning (PAL) researchers at the Air Force Research Laboratory (AFRL) are piloting an innovative integration of their cognitive modeling and personalized learning capabilities as part of the Air Education Training Command's (AETC) Linguist Next project. AFRL has teamed with the School for Standard Arabic at the Defense Language Institute (DLI) to provide optimized learning content and scheduling to both instructors and students, and are assessing the utility of this approach across multiple DLI training systems.

First, AFRL will embed their predictive analytics into DLI's forthcoming learning management system (LMS), Canvas. Canvas will provide detailed, fine-grained analysis of digitized student performance over time, and AFRL's predictive analytics will afford the opportunity to identify learning and forgetting rates associated with a depth and breadth of student competencies so that targeted refreshers may be presented to the learners at times deemed critical to help learners both acquire and sustain those knowledge components.

Second, predictive analytics may examine average learner profiles across the full Arabic curricula laid out within Canvas, to determine whether variations to that curriculum could better aid student acquisition and sustainment of skills. In this way, recommendations could be submitted to instructors and rolled out to future cohorts for validation and testing.

Third, predictive analytics will be integrated into a novel, virtual reality gaming environment for linguist learning under development by Jedburgh Technologies. This immersive learning application puts into context the vocabulary that learners must acquire in the course, and data are currently streaming in from DLI students for analysis. This technology is being used by students outside of classroom study and we aim to embed our predictive analytics into this platform in order to optimize and streamline student study time, and provide students with the right vocabulary content at the right time based upon their unique performance histories.

All of the aforementioned deliverables will be assessed across multiple cohorts moving through the 64 training curriculum at DLI. It is our 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, 711 HPW/RHWS

Photo by Mr. Will Graver

FIDELITY & READINESS ASSESSMENT METHODS FOR MEDICAL SIMULATION

Choosing the best simulator for training is challenging. Many considerations must be weighed by decision makers, including functionality, cost, reliability, maintenance requirements, flexibility, and capability for feedback. It is also important to consider the degree of realism or fidelity a simulator has compared to the real-life scenario it represents. For the military medical community, there is a paucity of objective data speaking to what characterizes a simulator as "high fidelity".

In recent years, the Personalized Learning & Readiness Sciences (PLRS) team has developed an integrative, conceptual model to represent the state of fidelity science. As part of this work, we have adapted a step-by-step method to standardize simulator fidelity assessment using empirical and self-reported data. Leveraging from our historical work in the context of flight simulation, we validated this method in a medical training contest.

For this work we selected surgical cricothyroidotomy, a high-consequence task currently simulated by technologies as diverse as sheep tracheas, augmented reality applications, and 3D printed skin. Validation involved conducting a comprehensive literature review to document the diversity of cricothyroidotomy simulators, collecting and analyzing performance data from experts and novices, and collecting questionnaire data from applied military purchasers and trainers. Analyses of these data have revealed valuable insights, including the perceived value of simulator fidelity for novices, experts and practitioners, the value of advanced data analyses to distinguish learner confidence from competence, and the capability of behavioral recording methods such as eye tracking and error analysis to capture learner intent, comprehension, and performance.

Task performance data demonstrated the utility of emphasis on transfer of training and error analyses. Self-report data demonstrated the importance of learner knowledge, confidence, and detailed judgments of simulator fidelity. Eye tracking metrics such as fixation counts and scan path efficiency ratings were shown to be sensitive reflectors of learner intent.

Findings from this research have been published across numerous conference proceedings (see below) and in a couple of forthcoming manuscripts. One recent manuscript focuses on learner calibration of competence and confidence. Combining performance and self-report data, the team leveraged novel statistical techniques to document low calibration by experienced task learners. A second manuscript (soon to be submitted for review) formalizes the steps taken to craft an integrated fidelity model. ☆

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

PUBLICATIONS

Bliss, J.P., Hartzler, B.M., Winner, J., & Hodge, D. (2020). Assessment and comparative evaluation of self-efficacy as a function of training style. *Proceedings of the 2020 Human Factors and Ergonomics Society Health Care Symposium (virtual conference).* Washington, D.C.: May 18-21.

Bliss, J.P., Etherton, K.C., Hodge, D., Winner, J., Millwater, T., McCool, J., & Cloude, E. (2021). *Applying error analysis to support surgical cricothyroidotomy simulator evaluation.* Poster presented at the 2021 Human Factors and Ergonomics Society Annual Meeting. Baltimore, MD: November.

Bliss, J.P., Etherton, K.C., Hodge, D., & Winner, J. (2022). Performance based evaluation of cricothyroidotomy simulators for training. *Proceedings of the 2021 Human Factors and Ergonomics Society Health Care Symposium.* New Orleans, LA: March 22.

Bliss, J.P., Etherton, K.C., Hodge, D., & Winner, J. (2022). *Using eye tracking metrics to quantify simulator learning by experienced and inexperienced medical trainees.* Presented at the 2022 Department of Defense Human Factors and Ergonomics Technical Advisory Group. Oklahoma City, OK: May 16.

Etherton, K.C., Bliss, J.P., Winner, J. (in preparation). *Examining Dunning-Kruger effects in medical training environments.* [Unpublished manuscript]

Etherton, K.C., Bliss, J.P., Winner, J. (in preparation). *Toward a singular model of fidelity: Incorporation of prior models within a novel framework.* [Unpublished manuscript].

SCALABLE TRAINING SOLUTIONS

Team Assessment in En Route Care and Beyond

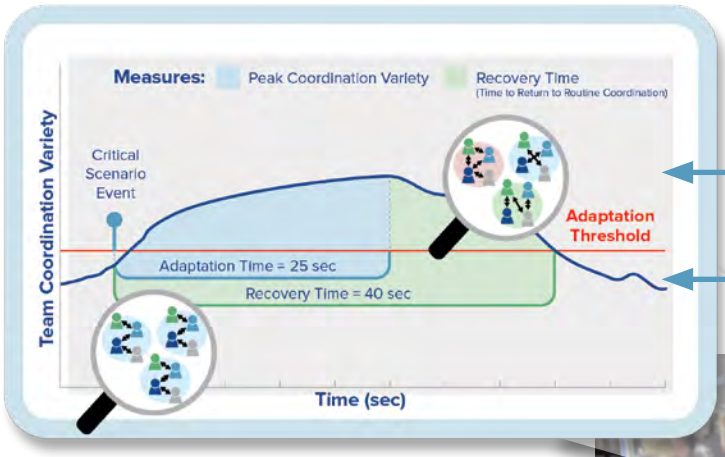
One of the top priorities within the Adaptive Proficiency Technologies line of effort is to define and objectively quantify readiness standards for individuals and teams. In real-world training environments, methods to quantify team coordination performance in complex mission scenarios are lacking. Our team is tackling these challenges in collaboration with the Georgia Institute of Technology, Arizona State University, and the University of Cincinnati. This work is leveraging over a decade of research within the team coordination space, enabling application of existing measures of team coordination to USAF use cases.

The first use case is Critical Care Air Transport (CCAT) teams. In partnership with the USAF School of Aerospace Medicine, we are demonstrating quantitative measures for team coordination behavior and developing real-time assessment tools to supplement the information available to instructors and learners. We are also working in coordination with Education Management Solutions to test prototype modifications to audio recording capabilities in simulation-based training scenarios. This work will enable scalable audio capture for heterogeneous training audiences and training testbeds.

Supplemental to the audio capture and coordination performance measures, our team is exploring solutions for integrating natural language processing to link to communication content, video analytics to extract scenario and performance ground truth data, and unobtrusive measures for identifying and predicting task saturation in individuals and teams. Taken together, these measures will feed into future use cases for providing sufficient discrimination between levels of achievement for team coordination within the context of military medical training.

Team coordination skill is one of the most universally applicable proficiencies in a training and mission performance context. This fact is driving our additional discussions about data formats and standards for how measurement and analytic tools built in the military medical context are applicable more broadly for other tactical training audiences. Investments in this area, initially focused on Aerospace and Operational Medicine, are poised to impact additional training audiences and operational contexts. ☆

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



Graphic by Ms. Shania Horner

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

RECENT PUBLICATIONS

Winner, J., King, J., Gorman, J., Grimm, D. (2022). Team coordination dynamics measurement in enroute care training: Defining requirements and usability study. In *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care.*

Grimm, D. A. P., Gorman, J. C., Robinson, F. E., & Winner, J. L. (2022). Measuring adaptive team coordination in an enroute care training scenario. Manuscript accepted for publication in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.*

Winner, J., Sorensen, D., Gorman, J. C., Grimm, D., Jarvis, M., Robinson, F., King, J., & DiBiase, R. (2022). Team coordination measures for en route care training: Foundations for scalable assessment. *Proceedings of the Military Health System Research Symposium.*

Photos by Mr. Gordon Klein

ADVANCED COMBAT TRAINING INNOVATION (ACTION) RANGE

The 351st Special Warfare Training Squadron (SWTS), home of the Pararescue and Combat Rescue Officer Apprentice Course, is tasked with leading, training, mentoring, and developing Air Force Special Warfare Operators in the art of Land Warfare and technical recovery skills to provide initial qualification training that meets warfighter requirements. Our training increases lethality, survivability, and readiness through state-of-the-art technology specifically designed to prepare warfighters to make discerning life-saving decisions. We need to develop tacticians, not simply marksmen.

In 2021, 351 SWTS built the Advanced Combat Training InnovatiON (ACTION) Range. This dynamic, live-fire adaptable training range uses innovative technology to provide analysis and feedback. "For decades, we have trained warfighters using predictable, immobile targetry to sharpen marksmanship skills. Although these skills are foundational, they do not drive real-world decision making and do not bridge the reality gap required to prepare warfighters for a lethal encounter." explained the Commandant of the Pararescue (PJ) Schoolhouse, Chief Michael Rubio.

The ACTION Range is the Air Force's first fully operational Shoot-Move-Communicate live-fire range. The concept is to provide combat-like situations in a controlled training environment. Ground combat fundamentals such as dynamic use of cover and concealment are rarely replicated in a training environment. In addition to providing a live-fire combat simulation, the ACTION range is also rich in data and video capture to allow for detailed after-action reviews (AARs) which reinforce student training and improve course delivery. In collaboration with RHWL, MVP Robotics developed Humanoid Engageable Kinetic Training Robot (HEKTR), which creates realism in training while capturing real-time objective data of performance.

> **"The ability to define product capabilities of industry to benefit the warfighter cannot be overstated. This is a unique opportunity to develop a product hand in hand with the company to meet training needs. Leveraging industry to close technical gaps is an operational imperative."**
>
> **— CMSgt Michael Rubio**
> Commandant of the Pararescue Schoolhouse

This prototype range and related technology have already paid dividends increasing both quality and time efficiency of training. The use of ballistically protected cameras for individual targets has reduced the weapon zero process by 75%, which allows shooters to make site corrections from their firing point without moving downrange. This feature alone increases efficiency while greatly mitigating risk. Additionally, the legacy weapons qualification course has been rendered almost obsolete by a more comprehensive evaluation of shooter proficiency. "We have not seen a single failure on the legacy Air Force Qualification Course (AFQC) for students completing the ACTION range Communities of Interest (COI).", said Chief Rubio. ☆

CMSgt Michael Rubio, Commandant of the Pararescue Schoolhouse

Maj Caitlin Harris, Special Warfare Prpgram Manager, 351st SWTS

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

Left: Rendering of PTRE site mapped with laser imaging, detection, and ranging (LiDAR) technology.
Top: Training exercise at the PTRE site.

Photo by Lt. Col. Deanna Bague
Rendering by Mr. Tyler Frost

PLAYAS TRAINING AND RESEARCH ENVIRONMENT

A top priority for Adaptive Proficiency Technologies is work to objectively quantify readiness standards for teams and teams-of-teams. Training infrastructure plays an increasingly important role in enabling scenario presentation, assessment, and feedback that scales to the team-of-teams end of the training spectrum. Bringing together disparate systems that are infrequently co-located requires careful planning to enable the digital training data required. This is true for scenario presentation, achieving appropriate levels of fidelity, and for enabling assessment and feedback.

RHW, along with the Information Warfare Division (RIG), the Layered Sensing Exploitation Division (RYA), and the Spectrum Warfare Division (RYW), are representing AFRL in a 2021 trilateral agreement with Air Combat Command (ACC) and Semi-Automated Forces (SAF) Concepts, Development, and Management Office. This team is partnered with New Mexico Tech to advance an experimentation site with training capabilities. Located in Playas, NM, the Playas Training and Research Environment (PTRE) site enables the integration of cyber and kinetic elements to create a multi-domain operations (MDO) capable training environment. A key goal is an integrated training environment for Electromagnetic Spectrum, Information and Cyber Physical Operations. Exposing multiple mission domains to an integrated live, virtual, and constructive (LVC) architecture is key to demonstrating MDOs and enabling mission outcome-focused training events in future contested and degraded environments.

The site is home to multiple Operational Test and Training research projects led by Mr. Ted Harmer, Research Electronics Engineer at RHWL, addressing cloud native

> **Our direct collaboration is critical for infrastructure development and implementation for PTRE to address current and future requirements for Operational Test and Training Infrastructures. We're engaging with our RHW Product Line Leads as well, which enables us to experiment with bringing in products that have grown out of RHW's 6.2 and 6.3 investments, such as the Performance Evaluation and Tracking System (PETS). These collaborations and this novel infrastructure and experimentation is expanding state-of-the-art and we're excited to see how this shapes OTTI and future LVC capabilities and what we're able to do along that path in collaboration with industry."**
>
> **— Mr. Ted Harmer**
> Electronics Engineer, 711 HPW/RHWL

Live Virtual Constructive (LVC) architectures (supporting Distributed Interactive Simulation protocols, IEEE 1278) and experimentation of LVC-enabled IoT devices. The data and results from the development and experimentation of technologies are immediately applied to the ACC training that takes place at PTRE, ranging from unit level to large force exercises. This enables an increase in data-driven debriefing, as well as powerful insights into technology applicability and feedback from operators early in the development lifecycle. ☆

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

HUMAN AND MACHINE CO-LEARNING

MISSION PLANNING AND DEBRIEF

The pointy spear tips of our national defense involve some of the most sophisticated technologies in the world. Just behind these spear tips, however, is a collection of planning and debriefing processes mired in decades-old technologies (e.g., chalkboards, whiteboards, laminated maps) that impede the operational ability to adapt, decide and execute. To address this, our team is leading research, development, and evaluation efforts for next generation mission planning and debrief. The goals are to streamline mission planning processes and promote human-AI collaboration to improve mission outcomes. This involves combinations of new technologies that allow human-AI teams to engage in these iterative processes together.

Our approach is to leverage state of the art Deep Reinforcement Learning (RL) algorithms to build computational agents that learn how to participate in the mission planning process by interacting with our custom learning environment. Metis, our in-house developed mission planning tool, provides the interface between computational agents and human mission planners. Our laboratory space at Wright-Patterson Air Force Base provides a testbed capability to evaluate the impact of collaborative AI on human performance in the mission planning task and to evaluate the performance of emerging AI technologies for collaborative planning.

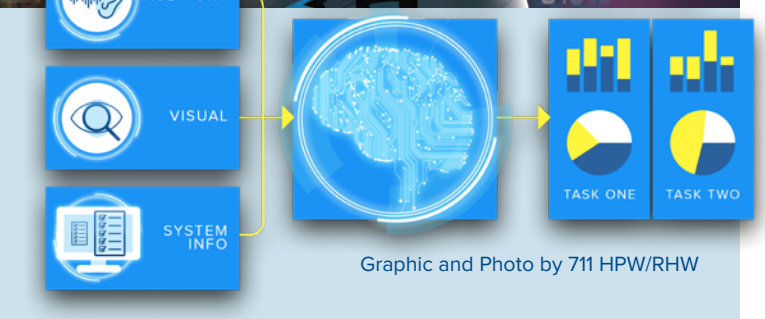
There have been significant advancements in theoretical and applied Deep RL in academia and industry within the past several years. However, a remaining significant challenge is how to design human-facing systems that utilize Deep RL in a way that promotes collaboration between humans and the computational agents. Effective human-AI collaboration requires that the complementary skills and abilities that both humans and AI possess must be fully leveraged. This will require coordination, cooperation, and communication between the human-user and agent. Our approach to automating the mission planning process incorporates these requirements with the goal of establishing a useful approach to building truly cooperative AI for a variety of DoD-relevant problems.

Photo by Mr. Will Graver

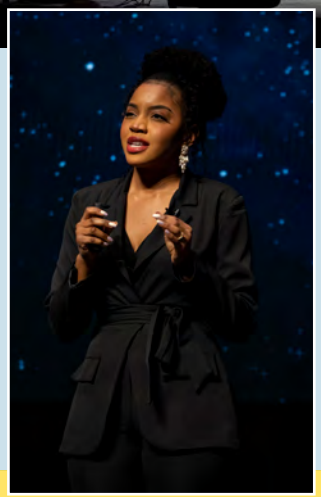
We are exploring extensions of the prototype capabilities described here. Our goal remains to push forward the state of the art of human-facing AI to support mission planning and leverage our knowledge and existing capabilities to move towards a broader set of functional mission types. ☆

Mr. Brandon Nolan, Program Technical Lead, 711 HPW/RHWL

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



Graphic and Photo by 711 HPW/RHW



Dr. Jayde King discusses futuristic scenarios of human-machine teaming during her AFRL Inspire talk at the Air Force Institute of Technology's Kenney Hall at Wright-Patterson Air Force Base, Aug. 23, 2022.

U.S. Air Force Photos
by Rick Eldridge

Watch Dr. King's AFRL Inspire Presentation HERE

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. Today, most autonomous agents are unable to perceive and anticipate user needs, goals, and status within a task. This lack of perceptibility 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 currently 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. We are now exploring methods to combine these two models while maintaining system transparency and explainability. The team is also investigating 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. ☆

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

VIA PUBLICATIONS AND PRESENTATIONS

Collins, G. M., Lee, M., Hough, A., Kennedy, S., & King, J. (July 2022) *Implementing online strategy and change detection algorithms in cognitive models using sequential Monte Carlo sampling for use in autonomous systems.* Presented at the Society of Mathematical Psychology, Toronto Canada.

King, J., Hough, A., & Collins, M. (October 2022) *A vision inference action model for shared situational awareness in autonomous agents: Approaches, challenges, & future directions.* Presented at the Human Factors & Ergonomics International Annual Meeting, Atlanta Georgia.

Montgomery, L. & Lee, M. (July 2022). *The wisdom of the crowd and framing effects in spatial knowledge.* Presented at the Annual Meeting of the Cognitive Science Society, 44. Toronto, Canada.

Montgomery, L. & Lee, M. (July 2022). *Applying cognitive models to wisdom of crowds with spatial knowledge extension.* Presented at In-Person MathPsych/ICCM 2022. mathpsych.org/presentation/795

THE MAKINGS OF A GOOD TEAMMATE

The Science of Understanding

"What makes a good teammate?"

The Science of Understanding team has been asking this exact question. With an eye toward designing machine technologies that will be good machine teammates, we must understand two things: (1) what characteristics or behaviors people think are important for other people to exhibit to be considered a good teammate, and (2) what might be unique or the same about the characteristics a machine must have to also be perceived as being a good teammate.

In a large-scale survey study, over 1200 owners and regular users of smart technology (all technology commercially available today, like GPS devices, smart watches and smart speakers) told us about how they use smart devices to help them out and rated the importance of 116 possible characteristics that could make either humans or those smart devices good teammates. The results show a strong consistency between perceptions of good human and machine teammates as shown in the table. However, because smart devices aren't yet capable of doing all these things very flexibly, on average, people rate all devices more like tools than like teammates.

Additional analyses are indicating that perceptions of the teammate-likeness of technology are influenced by individual variables, like age, amount of smart device use, and general tendency to anthropomorphize things in the world. Using novel mental model assessment tools, we are finding that people describe machine teammates with different terms than they describe human teammates. With these findings and insights in hand, the Science of Understanding team will now work with other AFRL efforts studying the collaboration and co-training between teams of humans and machines in order to test if designing highly rated and desired characteristics into new technologies can really strengthen human-machine teams, or to help both human and machine agents better understand each other's capabilities and limitations.

RECENT PUBLICATION

Blaha, L. M., Abrams, M., Bibyk, S., Bonial, C., Hartzler, B., Hsu, C. D., Khemlani, S., King, J., St. Amant, R., Trafton, J. G., & Wong, R. (2022). Understanding is a process. Frontiers in Systems Neuroscience (Special Topic on Understanding in the Human and the Machine), 16. https://doi.org/10.3389/fnsys.2022.800280

The designs for trusted, transparent, and effective teammates are enabled through the Science of Understanding. 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, Uninova Portugal, and the Pacific Northwest National Laboratory; it is the flagship research program for AFRL's Carnegie Mellon University Operating Location. ☆

Dr. Leslie Blaha, Senior Research Psychologist, 711 HPW/RHWL

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

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

Dr. Erin McCormick, Research Psychologist, 711 HPW/RHWL

		SMART DEVICE TEAMMATES	HUMAN TEAMMATES
#1	My teammate must be...	RELIABLE, COMPETENT, COMMUNICATIVE	COMMUNICATIVE
#2	My teammate must be capable of...	RELATIONSHIP BUILDING	RELATIONSHIP BUILDING
#3	To complete tasks, we must be able to...	DIVIDE AND CONQUER	DIVIDE AND CONQUER
#4	My teammate must be able to help me via...	COGNITIVE AIDING	PREDICTABLE AIDING
#5	My teammate must be...	TRANSPARENT	TIMELY
#6	My teammate and I need to have...	COMMON GOALS	
#7	My teammate must exhibit...	PREDICTABILITY	

GENERALIZED MODEL REPRESENTABILITY

ENABLING COMPARISON OF MAJOR THEORETICAL FRAMEWORKS IN SEQUENTIAL DECISION MAKING

To make the transition from research to operational human-machine systems, we need the ability to verify, validate and accredit cognitive models and artificial intelligence algorithms. However, we face a tall hurdle in achieving the desired evaluation and accreditation of formal models of intelligence: rapid advances in cognitive models and artificial intelligence are producing candidate models in many different formalisms, including statistical models, dynamic systems models, computational logic-based models, deep neural network models, quantum probability models, and more. Model comparison is frequently impeded by differences in the foundational assumptions between different formalisms. Comparison across formalisms is often impossible without numerous ancillary assumptions.

Researchers in the Human and Machine Co-Learning line of effort are working to answer the question, how can we evaluate, compare, and integrate theories developed in diverse modeling formalisms to bring us closer to a comprehensive understanding of human intelligence and cognition? To address this question, we introduced the Model Convergence Framework, shown in Figure 1. We reconceptualize model evaluation as requiring consideration of how similar candidate models are in their predictions about patterns of behaviors and in the cognitive mechanisms they use to explain the patterns. Using this framework, we can study models with many evaluation tools to identify which are offering convergence in mechanisms and predictions; convergence provides us confidence in our model evaluations.

To advance the science of comparison, we are also developing a generalized representation space. Each model can be mapped into this space, creating a potentially high-dimensional geometric surface. In such surfaces, we can measure properties of the models and directly compare them in ways not currently possible. We are demonstrating this evaluation technique on experiments in sequential decision making, where people are asked to make two related decisions. Models developed in the ACT-R computational architecture, in multinomial processing trees, and in quantum cognition formalisms all offer good candidate models for interesting patterns of sequential decisions. We are exploring how to use our new generalized representation space to bring these disparate approaches toward model convergence, seeking validated and ultimately accredited formal representations of intelligence for use in human-autonomy teaming technologies.

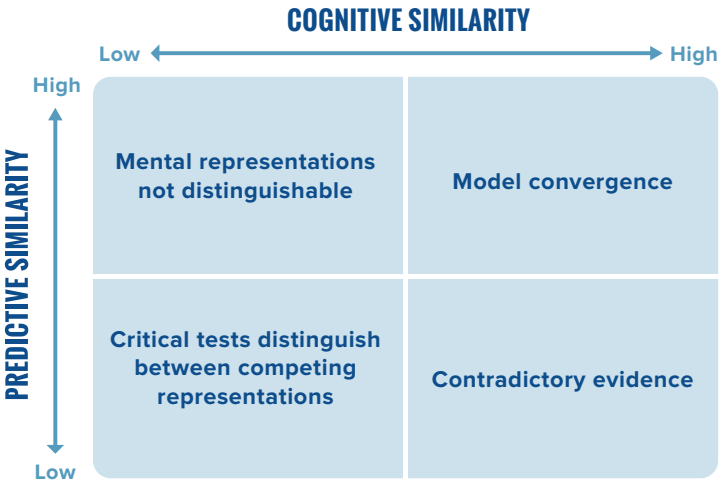


Figure 1. Conceptualization of the Model Convergence Framework

This research is supported by the Agile Science of Test & Evaluation portfolio of the Air Force Office of Scientific Research, Program Officer Dr. Brett Pokines. The vision of the Agile Science of Test & Evaluation is to advance foundations and capabilities of test, evaluation, validation, and verification in support of Department of the Air Force's test community needs. ☆

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

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

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

You can find out more about the Agile Science of T&E portfolio at:

https://community.apan.org/wg/afosr/p/agile_science_of_test_and_evaluation_tande

PUBLICATIONS

Fisher, C.R., Borghetti, L., Houpt, J.W., Blaha, L.M., & Stevens, C. (2022). A comparison of quantum and multinomial processing tree models of the interference effect. *Proceedings of the 20th International Conference on Cognitive Modeling.*

Borghetti, L., Fisher, C.R., Houpt, J.W., Blaha, L.M., Gunzelmann, G., & Stevens, C. (2022). Towards a method for evaluating convergence across modeling frameworks. *Proceedings of the 20th International Conference on Cognitive Modeling.*

Fisher, C. R., Houpt, J. W., & Gunzelmann, G. (2022). Fundamental tools for developing likelihood functions for ACT-R: A tutorial. *Journal of Mathematical Psychology, 107, 102636.*

Blaha, L. M. & Gluck, K. A. (in press). Computational model validation, comparison and selection. In R. Sun (ed.) *The Cambridge Handbook on Computational Cognitive Sciences.*

Fisher, C.R., Borghetti, L., Houpt, J.W., & Stevens, C.A. (2021). An ACT-R model of order effects. *Proceedings of the19th International Conference of Cognitive Modeling.*

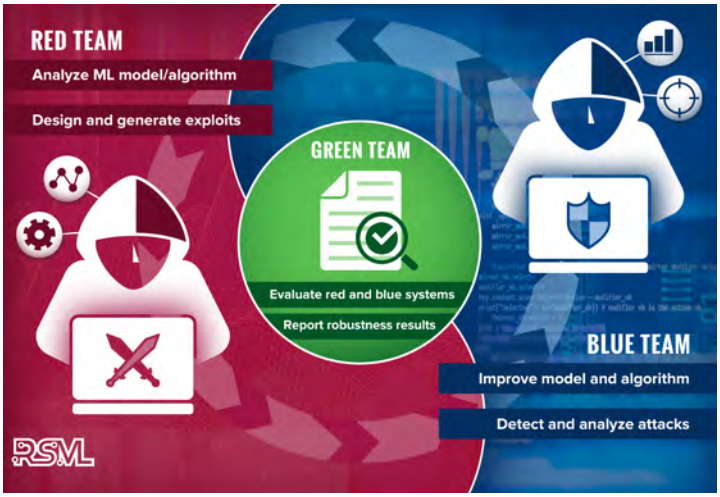
ROBUST AND SECURE MACHINE LEARNING

COMMANDER'S RESEARCH AND DEVELOPMENT FUNDS EFFORT

Machine learning image classifiers can be hacked. Several published demonstrations have found ways to add small changes to images or changes to the algorithm training processes that cause the resulting classifier to make errors. Popular examples include making machines label turtles as rifles and cats as guacamole. The U.S. Air Force and U.S. Space Force are enthusiastic about finding ways to leverage machine learning (ML) and artificial intelligence (AI) to improve warfighter capabilities. We must have a deep understanding of the ways these emergent technologies are vulnerable to adversarial attacks before we put them to work in our missions.

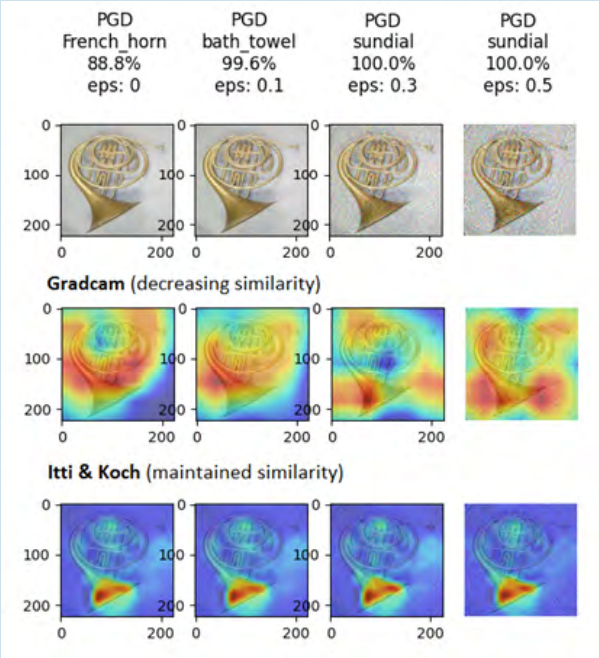
To meet the call for investments in "resilient, robust, reliable and secure" AI/ML technologies (2018 DoD Artificial Intelligence Strategy), researchers from the Airman Systems Directorate, led by RHW, have partnered with researchers from the Information, Sensors, and Munitions Directorates in an AFRL-wide Commander's Research and Development Funds effort investing in Robust and Secure ML. Since this research and development (R&D) commenced in 2020, this team has demonstrated a new approach to evaluating adversarial attacks by adopting a Red-Blue-Green teaming approach that divides responsibilities across Red Team offensive attacks, Blue Team defensive models, and Green Team objective evaluation. RH has a strong track record in conceptualizing and quantifying robustness, our RHW team serves predominantly in the Green Team role.

A key outcome of the work is the implementation of the Adversarial Robustness Assessment Toolbox (AdversarialRAT), a suite of tools for systematically testing image classifiers and measuring robustness to many types of adversarial ML attacks. We are demonstrating the capabilities of the AdversarialRAT in a series of challenge events pitting the Red and Blue Team technologies against each other. We have successfully deployed multi-dimensional assessments of robustness, including the first test of a novel metric, the Index of Robustness. The challenge events firmly established a baseline for existing ML capabilities and identified critical directions for the research into AI/ML robustness to meet unique Air Force and Space Force needs.



Graphic by Ms. Shania Horner

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.



The top row shows an image under an increasingly strong (from left to right) adversarial attack; readers can see the changing label that the machine would apply to the French horn. The middle row illustrates how the ML algorithm's perception changes under attack, and the bottom row shows how human perception is not strongly influenced by the image attack.

In 2022, our team has continued pushing the state of the art in quantifying classifier performance. We have developed new methods for assessing how adversarial attacks differently impact human and machine perception, because the difference between human and machine capabilities may suggest ways that teaming them together may improve robustness. We are currently testing the ways that humans may contribute to the robustness of attack detection through active learning. Teaming through co-learning between humans and machines to create robust classifiers is an exciting direction for future capabilities. ☆

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

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

ADAPTIVE WARFIGHTER INTERFACES

CORE TECHNICAL COMPETENCY (CTC)

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

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 artificial technology (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 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.

Continued on next page

Photo by AFRL

FUTURE WARFIGHTING ENVIRONMENTS WILL COMPLICATE DECISION-MAKING IN SEVERAL WAYS:

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

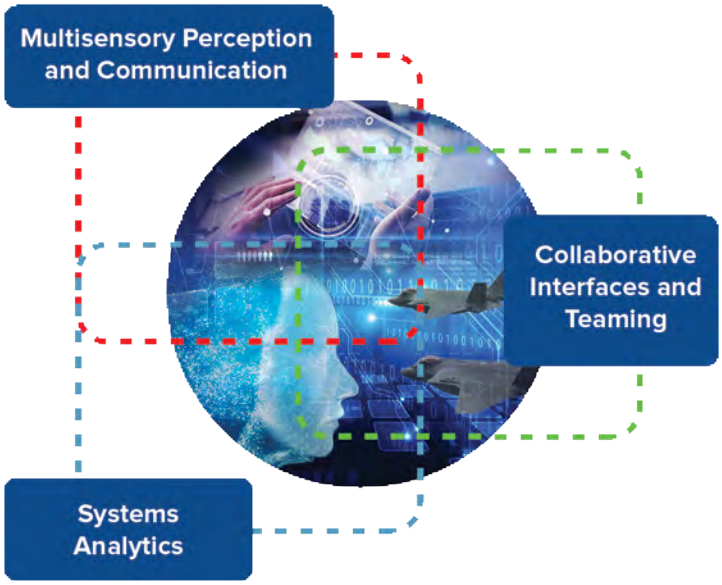
Fallible machines that continually gain intelligence, learn, and adapt over time

High-stress environments requiring rapid, accurate situation assessment and 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 core research areas (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 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.



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.

Continued on next page

THE AWI RESEARCH PORTFOLIO FOCUSES ON THE FOLLOWING FUTURE CAPABILITY VISIONS:

Joint All-Domain Command and Control (C2)

Manned-unmanned teaming

Data-driven Intelligence, Surveillance and Reconnaissance (ISR) fusion

Emerging Space applications of warfighter interfaces

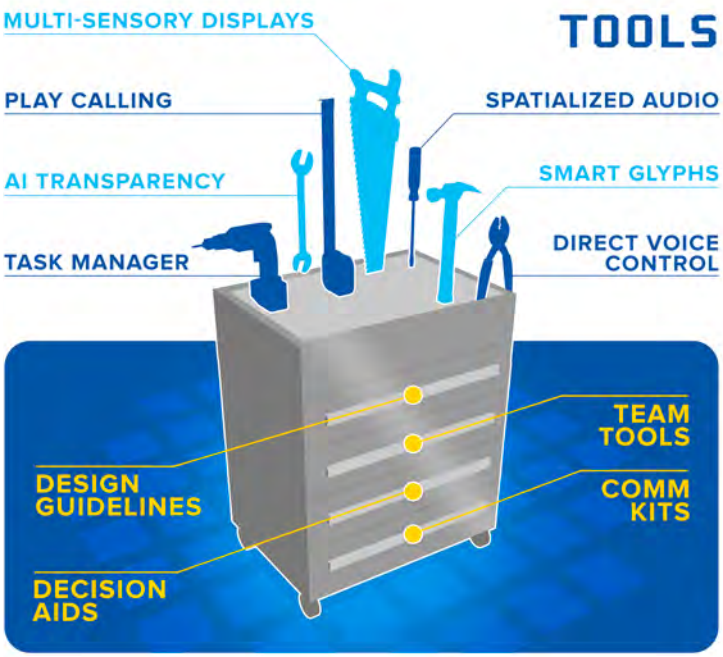
The AWI CTC is executing specialized, warfighter-centric interface research that addresses high priority USAF needs and optimizes warfighter decision making."

—Dr. Mark Draper,
Adaptive Warfighter Interfaces CTC Lead

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.

The AWI CTC is executing specialized, warfighter-centric interface research that addresses high priority USAF needs and optimizes warfighter decision making. With projects focused on improving warfighter performance in JADC2, Space, Electronic Warfare (EW), ISR, and Cyber environments, FY23 looks to be another exciting and productive year! ☆

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



Graphic by Mr. Will Graver

AWI ACCOMPLISHMENTS:

- **Initiated new Joint All-Domain Planner with Adaptive Collaborative/Control Technologies (JADPACT) advanced 6.2 research project. JADPACT will design, develop, and evaluate advanced warfighter interface & decision-aiding technology concepts for a Joint All-Domain Mission Commander Control Station, enabling the effective C2 execution of integrated all-domain effects.**
- **Graduated core technology into 6.3 Program that will increase speed of confident decision making by enabling identification of intel gaps through structured analytical techniques, automating collection requirements, and optimizing warfighter ISR planning efforts (strategic to tactical).**
- **Established seat-hopping testbed with AF Test Pilot School and Univ. Iowa for rapid flight testing of novel interfaces that accelerate and extend a warfighter's situational awareness (SA) and decision making with multiple Unmanned Aircraft System (UAS).**
- **Designed and implemented plug-in framework for training, visualizing, and fielding eXplainable AI models. Automated AI/machine learning (ML) tools and visualizations to support flight path planning using multidimensional data.**
- **Completed detailed Cognitive Task Analyses with warfighter experts to identify key information requirements associated with:**
  - **Maintaining pilot SA with autonomous wingman**
  - **Developing a conversational AI agent for Intel Analysts**
  - **Improving satellite operations**
- **Completed Field Assessment of F-35 Automatic Ground Collision Avoidance System (Auto-GCAS) New Air Force Office of Scientific Research (AFOSR)-funded foundational research: "Positively Influencing Attitudes and Behaviors Through Social Media Messaging."**
- **Initiated new seedling research efforts in several areas including:**
  - **Exploring adaptive strategic reorientation of attention**
  - **Holding algorithms accountable: Detecting implicit bias in Machine Learning annotators**
  - **Low Frequency/Infrasonic Data Communication Profile**

SYSTEMS ANALYTICS (SA)

CORE RESEARCH AREA

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

TWO PROJECT AREAS

- **Analytic-enabled Cognition**
- **Sense-making at Scale**

Work domain analyses and technology surveys reveal that contemporary products frequently used by our warfighters are not designed to dynamically integrate machine analytics and human cognition into an effective joint cognitive system. Few fielded capabilities exist that promote sensemaking of complex, uncertain, and multi-dimensional data, and most that do lack the dynamic tailoring of their processes to meet situation-specific warfighter needs. As a result, operators across mission areas are challenged to rapidly make meaning from "Big Data" and must overcome the inherent complexity and uncertainty of multi-dimensional data and associated analytics to create actionable insights in real time.

To mitigate this shortcoming, research within Systems Analytics focuses directly on the discovery and development of Airman- and Guardian-centric tools and solutions that enhance the cognitive impact of analytics on thinking and reasoning: the synthesis of context-specific machine analytics, algorithms, interfaces, and advanced visualizations that improves performance by creating integrated joint cognitive systems to reduce cognitive burden and do work on behalf of the warfighter.

The components of the Systems Analytics portfolio exist to assess and enhance the impact of analytics on thinking and reasoning in order to tailor capabilities to the context-specific cognitive requirements of our warfighters. The result is the design and creation of analytic solutions (i.e., analytic products, interfaces/visualizations, publications, and guidelines) that enable superior decision-making and enhance rapid sensemaking to directly support warfighter needs.

Continued on next page

Graphic by 711 HPW/RHWA

SYSTEMS ANALYTICS (CONTINUED)

Specific tactical pivots and optimizations are already in progress, and further increase the value of the Systems Analytics portfolio:

- Ensure strategic alignment with DoD, AF, and AFRL priorities, such as deliberate connections to Operational Imperatives, Joint All-Domain Command and Control (JADC2), Science and Technology (S&T) 2030 objectives, and similar high-priority investment areas

- Concentrate on high-value enduring research that can be transitioned to Airmen and Guardians: rapid sensemaking for decision superiority; data integration, aggregation, and visualization for Intelligence, Surveillance and Reconnaissance (ISR); artificial intelligence/machine learning (AI/ML) capabilities to mitigate "Big Data" complexities

- Focus on development of publications and unifying Design Guidelines to inform appropriate design and evaluation of joint cognitive systems across USAF/USSF/DoD

- Build and strengthen coalitions of expertise across AFRL, USAF, DoD, academia, commercial, and international partners (especially leveraging linkages to the United States Air Force School of Aerospace Medicine (USAFSAM), per RH Director's Intent)

- Increase the use of modeling and simulation as a tool for high-risk/high-reward research

- Embrace the dichotomy of meeting urgent (short-term) operational needs while simultaneously ensuring the DoD is well-positioned to win the (long-term) future fight

The Systems Analytics portfolio is versatile and robust, providing the Air and Space Forces and the DoD strong analytic tools and decision aids, as well as high-impact and timely deliverables. ☆

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

ANALYTIC-ENABLED COGNITION

Many current machine learning and artificial intelligence approaches employ analytic-based solutions that manipulate and present 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. Research in the Analytic-enabled Cognition line of effort fills that gap, specializing in providing Airmen and Guardians with the foundations of analytics and visualizations deliberately designed to enhance their decision-making needs and help them rapidly make meaning from complex and uncertain information.

A variety of novel analytics studies and research activities are underway, listed here with expanded details in the sections that follow. This year, a new analytic designed to support and enhance the performance of intelligence analysts was started ("Cognitive and Physiologic Performance"), with two research studies planned for the first half of FY23. The Systems Analytics Assessment team continues to investigate which human-machine systems improve the meaning-making capability of intelligence analysts, with publication of results and new study design in progress. Several journal articles were published to capture results of studies and methodologies, including work investigating the impact of human-machine interfaces and decision analytics, with additional articles being developed. Significant work continues in meaning-making research, developing guidelines and methodologies to help recognize and mitigate adversarial influence. This research also focuses on recognition, capture, representation, and responses for adapting to social-cyber behaviors and providing resilience to misinformation. Work is also ramping up in the development and evaluation of natural language interfaces and analytics for human-machine teaming solutions in the cross-directorate Conversational AI Assessment research. ☆

Dr. Vincent Schmidt, 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 is designed to test the viability of physiological measures to provide real-time information to supervisors of intelligence analysts to maximize their performance and health. Based on knowledge extraction interviews and a Cognitive Task Analysis of Air Force analysts, an unclassified simulation of the analyst tasks, the Joint Activity Design and Evaluation (JADE) platform, has been developed to support data collection from current Air Force analysts. JADE simulates the typical interactions of individual analysts with representative data types such as maps, photos, and human intelligence reports in order to solve scenario problems. Although the scenarios are fictional, the problems posed and the available data simulate those confronted by analysts in the typical work shift. Scenarios are designed to inflict different demand levels by varying the amount of information and the degree of agreement between different pieces of information. The scenarios were developed by an expert analyst and underwent extensive validation by additional analyst subject matter experts.

STUDY #1:

Will test the sensitivity of both the performance and the physiological metrics to manipulations of scenario task difficulty.

STUDY #2:

With the addition of data collection before and after the analysts perform a duty shift, this study will enable a direct test of the sensitivity of the physiological measures to the fatigue induced by a full shift's work.

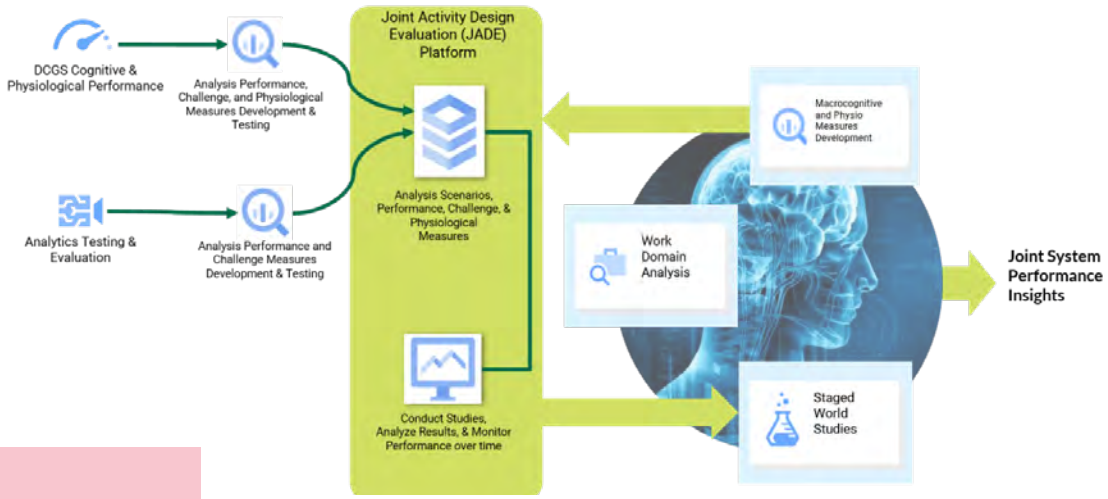
Physiological measures will be collected as analysts perform these scenarios. The physiological measures have been selected to represent those that combine the most robust assessment of human cognitive state along with potential unobtrusiveness in operational settings. The measures include various heart-and eye-based measures.

The first study is currently scheduled to begin data collection in October 2022 at Hurlburt Field. It will test the sensitivity of both the performance and the physiological metrics to manipulations of scenario task difficulty. Since the physiological measures provide continuous data streams, the first study will provide the opportunity to determine whether the measures could provide analyst supervisors with indications of which team measures were contending with potential data or mental workload overloads. Such detections could enable the supervisor to take actions such as shifting assignments to aid the overloaded analyst.

The second study is planned to begin data collection within a few months of the first study. The design will be very similar, with the addition of data collection before and after the analysts perform a duty shift. This will enable a direct test of the sensitivity of the physiological measures to the fatigue induced by a full shift's work. Along with the performance measures of how well the analysts perform in the JADE scenarios, the second study will test how well the physiological measures might help supervisors of analyst team identify team members potential at risk of sub-par performance.

Overall, CPP's use of carefully crafted simulations and scenarios with operational Air Force analysts as participants will determine the viability of physiological measures for improving analyst operations. ☆

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



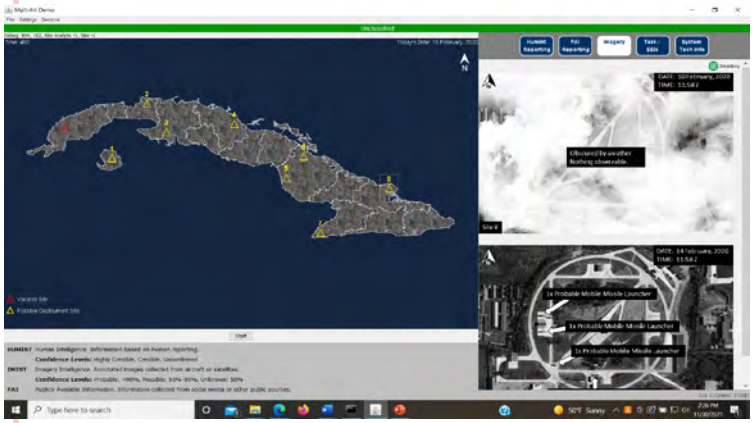
Graphic by Mile 2

Screengrabs 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 to predict and anticipate future events. One way to enhance meaning-making may involve collaboration from a machine agent, such as one that uses artificial intelligence (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 recognition aid that highlights persons and vehicles (similar to Project Maven) or by a simulated agent that directed attention to EEIs. Full motion videos of simulated compounds as viewed by an overhead camera were created. Each video had a specific conclusory event and an embedded sign that would serve as an indicator of the event earlier in the video. Novices and expert intelligence analysts participated in the study. Results showed that EEI identification was enhanced by having the simulated agent direct the participants' attention to the EEIs, but this recognition aid did not improve prediction of the event nor did it enhance the ability to recursively identify the sign that could have helped predict the event. The recognition aid that highlighted persons and vehicles did not improve EEI identification, but it did improve meaning-making for predicting events when the sign was a person or a vehicle. It did not, however, enhance meaning-making when the sign was not a person or vehicle. This implies that neither recognition aid served to free the cognitive resources of participants to better make meaning, but that meaning was enhanced as a function of drawing attention to the appropriate signs. 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 goal of Multiple Intelligence (Multi-INT) Analytics research is to understand how adding decision analytics to a task in which people must combine different types of information from different sources affects their ability to make a cohesive assessment of the situation. The first study looked at how analytic accuracy and transparency affected performance. People were given access to Publicly Available Information (PAI) data (Twitter feeds, Facebook posts), overhead imagery, and human intelligence (HUMINT) reporting before being asked to determine the most likely site of an Integrated Air Defense System. Results revealed that performance (accuracy) in the high accuracy condition was significantly lower when the analytic was transparent than when there was no transparency. For the analytic with low accuracy, transparency made no difference in performance. The next study will further explore conditions under which transparency is beneficial for performance and when it is not. ☆

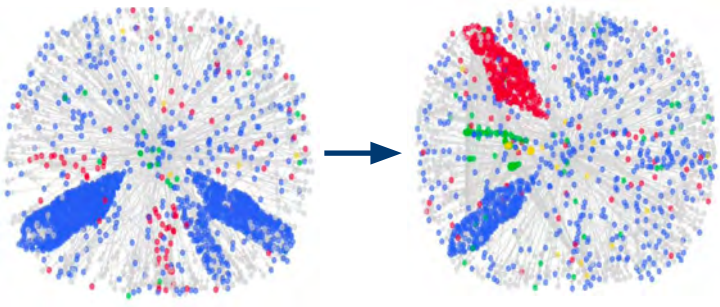
Dr. Kellie Turner, Research Psychologist, 711 HPW/RHWAR

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

SUB-TASK 2.1: INFLUENCE CHARACTERIZATION

A NATO Review article recently stated that social media, social-cyber behavior, and mobile technology is enabling a new domain: Cognitive Warfare. The goal of cognitive warfare is to influence what people think and how they act by shaping individual and group beliefs and behaviors in ways that favor adversarial objectives. Without engaging in force or coercion, deleterious outcomes can occur such as fracturing institutions, groups and even an entire society. Previously, many case studies regarding malign influence involved information maneuvers involving text content (e.g., Facebook or Twitter). Current trends in social-cyber behavior show increasing use of different platforms (e.g., TikTok, WeChat) and use of both audio and visual content (videos, memes, etc.). In a keynote speech recently given in a Political Memes Workshop in the 2022 International *Association for the Advancement of Artificial Intelligence* (*AAAI)* Conference on Web and Social Media (*ICWSM*) Dr. Kevin Munger stated emphatically that the analytic approaches developed to understand influence impacts based on text will not work with visual content. MMIE and others, having sensed this shift are increasing multi-platform and visual content influence characterization and assessment research.

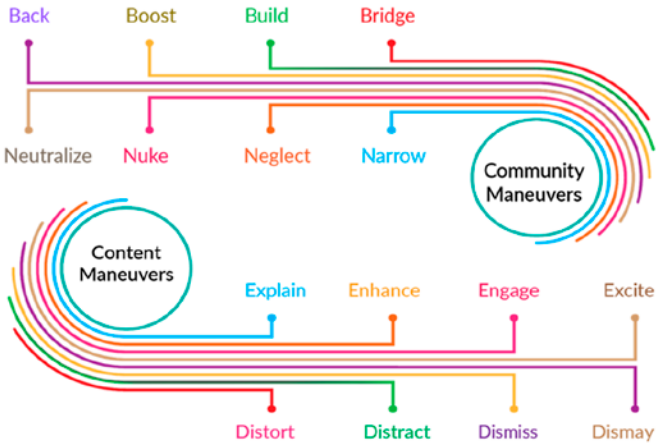
Anti-vaccine community one week later



Graphic by Dr. Nitin Agarwa, University Arkansas Little Rock (UALR)

University of Arkansas at Little Rock has pioneered influence characterization analytics for visual content. One recent finding is that increasing toxicity in content foreshadows splintering in communities (and increasing polarization). Community splintering and polarization are exacerbated by platform algorithmic biases relative to context and content, which creates online extremist content communities and boosting of artificial engagement (bots, etc.) In addition, techniques have been developed to characterize multi-platform influence (e.g., tweets that contain links to YouTube videos, etc.) Two Small Business Innovative Research (SBIR) efforts will begin this fall that will develop, demonstrate, and transition analytics to detect, classify, and forecast the impact of information maneuvers based on visual content (video, memes, and pictures) and audio chat. The SBIRs utilize the

BEND Information Maneuvers



Graphic by Dr. Rebecca Goolsby, Office of Naval Research (ONR)

BEND framework, developed by Dr. Kathleen Carley of Carnegie Mellon University, MU) which identifies information maneuvers influencing both networks and content/narratives. The SBIRs will support meaning making and decision-making regarding influence, course of action assessment and forecasting of behaviors, events.

Laurie Fenstermacher and Kirsten Rice are engaged in defining requirements and research gaps for a Social Digital Media Playbook to support Information Operations and Public Affairs operators, under a Headquarters Air Force/ Deputy Chief of Staff for Operations (HAF A3) Operations in Information Environment (OIE) Strategic Plan. In addition, Ms. Fenstermacher and Ms. Rice are analyzing existing frameworks and developing new frameworks that identify information maneuvers influencing both networks and content/narratives. New work is focused on adding to both the list of maneuvers based on new platforms and new types of content (visual, audio), as well as developing Human Machine Interface concepts for enabling operators to make meaning based on tailored information/graphics.

In a program focused on the development of anticipatory analytics for Air Mobility Command missions, AFRL conducted a case study to demonstrate text analytics and forecasting modeling capabilities on Ukraine (with Twitter and Reddit social media data and events data). As in previous forecasting studies, many significant indicators of Russian hostility were identified, including sentiment, psycholinguistic variables such as cognitive complexity and idea density, language associated with urgency, and polarizing language. The anticipatory text analytic toolsuite, SAEText, has been transitioned to 618 Air Operations Center, Strategy Directorate (AOC/SRD) in Air Mobility Command. ☆

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

SUB-TASK 2.2: RESILIENCE TO MISINFORMATION

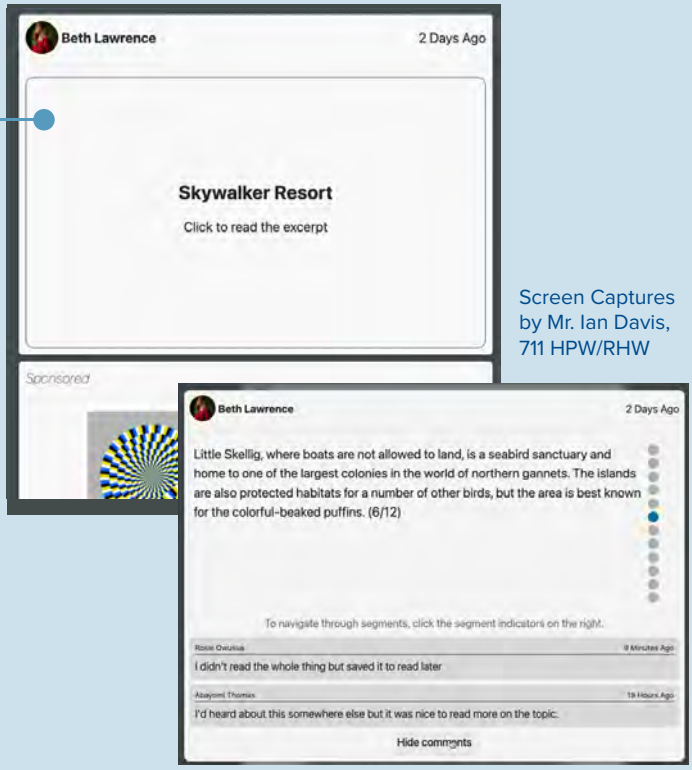
In addition to the research on social media analytics, the MMIE portfolio includes behavioral science efforts to examine the underlying mechanisms of influence. Currently, there are four studies under this objective: (1) Bias in artificial intelligence/machine learning (AI/ML) annotators, (2) Human perception of auditory deepfakes, (3) Gamification of insider threat communication indicators, and (4) Narrative framing and individual differences in a social media emulator.

The social media emulator project, led by Dr. Kathleen Larson, has progressed substantially this year, with support from Dr. Megan Morris (RHWM) and Dr. Christine Vitiello as co-PIs, and CAE and Ball Aerospace as contractors. The team is conducting research to support Operations in the Information Environment (OIE), which is defined as the sequence of activities that use information to affect behavior and to influence actors, information networks, and systems (Headquarters Air Force, 2020). Therefore, a critical next step in this area is understanding how the presentation of information influences memory for the content.

In a proof-of-concept experiment, a lightweight social media emulator was developed to assess the influence of text framing on what information users retain. Participants were recruited online through Amazon's Mechanical Turk and then read several text excerpts either in an emulated social media environment (resembling Facebook) or a traditional web-based survey format with the content presented as fiction, nonfiction, or fake news. For example, in the fake news condition, participants were told to read excerpts as they would normally read a social media post. After reading the excerpts, participants then took a recognition test to determine to what extent they were remembering the content accurately. Additional measures were included to examine the influence of latent individual differences (e.g., Big 5 personality traits) on memory performance.

Preliminary analyses suggest the different framing conditions did not influence memory. However, in some cases, memory performance was weaker when excerpts were read in the emulator compared to the traditional format. These results will be shared at a meeting for the Society for Computing in Psychology (SCIP) in November, along with the individual differences data, and lessons learned about conducting research in a social media emulator. An extension of this work will be executed under a Laboratory Research Initiation Request (LRIR) in 2023 to examine factors related to the impact of social media messaging and how it influences online to offline behavior. Overall, this research is the first step in developing a virtual social-cyber testbed to elucidate how psychological processes interact with different media, formats, and algorithms in the Information Environment. ☆

Dr. Kathleen (Katie) Larson, Research Psychologist, 711 HPW/RHWAR



Screen Captures by Mr. Ian Davis, 711 HPW/RHW

Top: Screen capture of the social media emulator feed.

Bottom: Users could click to read full text excerpt and to see comments.

TASK 3: CONVERSATIONAL AI ASSESSMENT

The Conversational Artificial Intelligence (CAI) assessment research includes analytics that translate user verbal or textual input into actionable machine analytics, giving a user some type of output. The Information Directorate is developing CAI that will integrate with complex analytics to provide input for analysts when solving Requests for Information. These capabilities will only be effective if the types of queries it supports are intuitive and relevant to analyst information needs. To identify and understand this requirement, our project team conducted a series of interviews with analysts representing various specialties and mission sets within the Intelligence Community (IC). A total of 12 analysts were presented with a fictional but realistic and representative mission scenario, developed by IC subject matter experts. Aggregated results from these interviews indicate that most queries will likely be associated with research and analysis. This effort will ultimately yield design and functionality requirements for inserting CAI technologies into analyst workflows across the IC. ☆

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

SENSE-MAKING AT SCALE

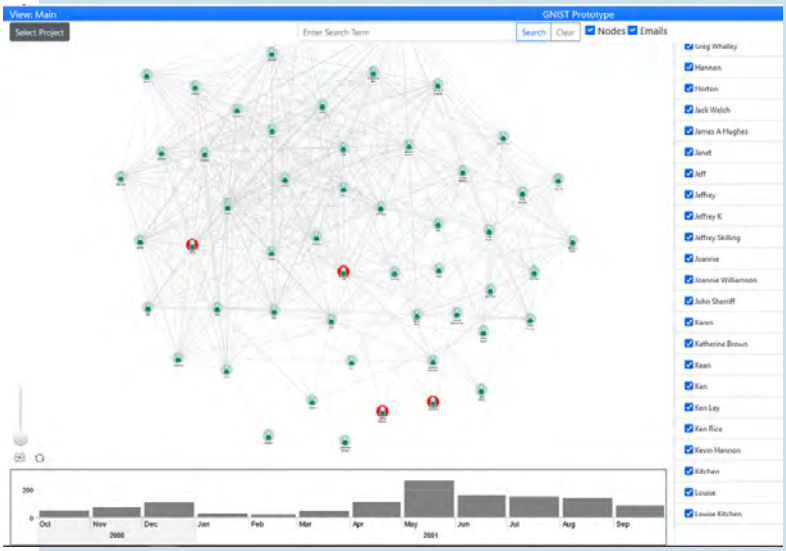
The Sensemaking at Scale (SaS) line of effort concentrates on research designed to enhance sensemaking and mitigate data overload, enabling Airmen and Guardians to rapidly extract meaning from complex, uncertain, and multi-dimensional data sources. The work in this portion of the portfolio focuses on developing specialized and tailored solutions to improve rapid sensemaking of “Big Data” in the face of complex, uncertain, and multi-dimensional / multi-sensor datasets. Much of the work in this group is directed toward specific and often urgent DoD and USAF customer needs.

The SaS collection primarily includes mid-late 6.2 work, including prototypes, operational demonstrations, customer evaluations, and in-progress product deployments. Primary projects are listed here and detailed below. The Generalized Network Insights from Situated Text (GNIST) analytic uses deliberate applications of graph theory to visualize rapid calibrated sensemaking. Work by the eXplainable Artificial Intelligence AI Lab (XAIL) has resulted in a software framework for the rapid exploratory data analysis and visualization, and the XAIL group is working with several operational and research groups to test and deploy its product. The Human Language Technology team has a wide range of products and corresponding research in transition, some supporting urgent DoD needs, and some feeding AFRL research tasks. The Dynamic Wide Area Design and Exploration (DWADE) research is engaged in several advanced studies, and the Joint All-Domain Integrated Intelligence, Surveillance and Reconnaissance (ISR) (JADII) research team is finalizing work on ISR optimization. ☆

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

The Sense-making at Scale line of effort concentrates on research designed to enhance sensemaking and mitigate data overload, enabling Airmen and Guardians to rapidly extract meaning from complex, uncertain, and multi-dimensional data sources."

— Dr. Vince Schmidt
Systems Analytics Core Research Area Lead, 711 HPW RHWA



Screen Capture by 711 HPW/RHWAR

TASK 4: APPLIED EXPLAINABLE ANALYTICS (AXA)

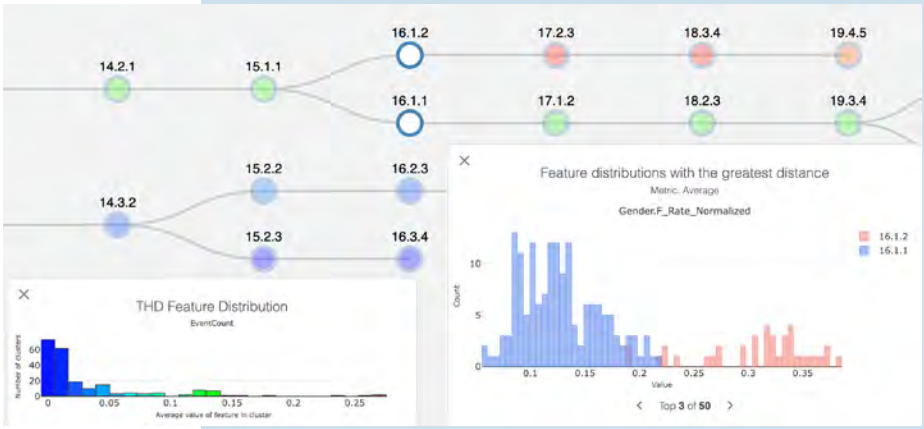
SUB-TASK 4.1: OPERATIONAL ANALYTICS

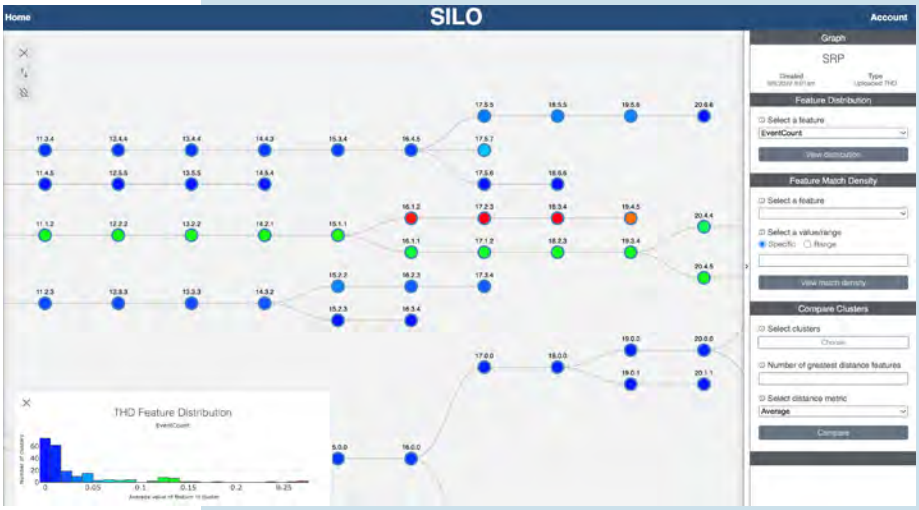
Generalized Network Insights from Situated Text (GNIST) looks at how a combination of analytics and user-centered visualizations can best enable rapid and calibrated characterization and sense making of complex networks. GNIST operates on unstructured text to discover relationships and is capable of handling 50+ languages to perform Natural Language Processing (NLP), including topic modeling and named entity recognition. Its current use case focuses on sensemaking of large email data sets for use by the Department of the Air Force Office of Special Investigations (DAF OSI). Over the next year, there are plans to install it on a stand-alone computer at DAF OSI so they can begin using it with their data and provide feedback. In conjunction, research is in the planning stages to investigate how GNIST facilitates understanding in these large data sets and the type of use cases for which it is most beneficial. ☆

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

SUB-TASK 4.2: APPLIED TOPOLOGICAL DATA ANALYSIS (TDA)

Screen Captures by 711 HPW/RHWAI





XAIL research includes several components:

- Fundamental evaluation and analysis of data types and datasets – research focused on understanding the suitability of TDA techniques and specific machine learning approaches to analytical cognitive work performed by our Airmen and Guardians. These studies started with XAI engagement through the Commander's Research and Development Funds.
- Framework development – the implementation of SILO, a novel lightweight and modular software system and computing architecture capable of ingesting a variety of data, applying selected machine learning approaches, generating visualizations, and exporting the results. SILO is powered by custom analytic engines developed by the XAIL team, and is designed as a platform supporting rapid data exploration, feature engineering, and customizable visualization. The SILO platform design and implementation is actively being pursued and extended to include more analytical techniques.
- Applied research – XAI expertise leveraging SILO and its components offer tailored machine analytics and explainable visualizations, providing insights into complex and multidimensional datasets. The XAI Lab's machine learning tools, including SILO, have successfully generated insightful visualizations for tasks such as Suicide Risk Prevention data and predictions of success for training candidates in certain programs.

Topological Data Analysis (TDA) is a subset of eXplainable Artificial Intelligence (XAI) with a strong basis in mathematics, computational analytics, and graph theory. TDA is one of the machine learning techniques leveraged by the XAI Laboratory (XAIL) team to perform advanced exploratory data analysis and knowledge discovery.

XAIL is actively connecting to critical operational USAF needs. The team is starting new collaborative work with the United States Air Force School of Aerospace Medicine (USAFSAM) this year. SILO will be extended to analyze biometric and training data, predicting and visualizing selection and training success metrics for USAF pilot candidates. The resulting models are expected to dramatically reduce training costs and program risks, saving time and improving the quality of candidates and applied training. The XAI Laboratory is also working with the Raptor/Lightning (F22/F35) program offices to use XAI models for enhanced flight path planning.

XAIL is successfully integrating research to solve some of the most important problems within the broader Air Force Community. XAI research is about making artificial intelligence explainable, improving the capabilities of the human/machine team by enhancing rapid understanding of complex data to improve meaning making, reduce complexity, and enable the warfighter to make better, faster decisions. ☆

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

TASK 5: HUMAN LANGUAGE TECHNOLOGY (HLT)

The Human Language Technology (HLT) Group of RHWAI successfully adapted and integrated a custom, license free, open-source Automatic Speech Recognition (ASR) system for the Network Integrated Training Environment (NITE) lab. Customization of the general model tuned towards the military domain of interest more than halved the measured error rate and still provided for near real-time transcription of the exercises being conducted.

Additionally, the HLT Group was able to showcase their ability for rapid response when operators reached back to AFRL for resources in Ukrainian ASR and Machine Translation (MT). The HLT's Ukrainian ASR system was requested and fielded in less than a week to an operational speech-to-speech translation system, and both ASR and MT models were fielded in less than a month to other partners assisting in the current conflict in Europe.

Previous efforts in Text To Speech (TTS) synthesis have been invigorated with the exploration of Neural TTS methods. The expertise and computational capacity of the HLT Group has combined with the Behavioral Scientists of RHWAR for Research and Development of Auditory Deepfake creation and an ongoing experiment to test resilience to disinformation. Related to this work is other research into how and with what words synthetic agents might actually communicate. Five summer interns were hosted and put forth a concerted effort toward building a Tailored Expressive Dialogue Systems (TEDS) based around chatbots and tailoring their output.

As an expansion to the current expertise of the HLT Group, methods in Multimodal Learning are being explored on multiple fronts. One researcher is investigating a technique to use objects in video scenes to improve machine translation of spoken dialog by providing necessary context. Another researcher is exploring Optical Character Recognition based methods to supplement baseline text translation and how this could lead to more resilient and robust Multimodal Translation systems.

Finally, the group is leveraging technology received from other agencies, such as the Defense Advanced Research Projects Agency (DARPA) and the Intelligence Advanced Research Projects Activity (IARPA), along with in-house developed capabilities, and has integrated and fielded this technology into a Cross Lingual Information Retrieval system. Further enhancements to data visualization and algorithmic robustness within the application is planned for the coming year. ☆

Mr. Eric Hansen, Senior Electronics Engineer, 711 HPW/RHWAI

RECENT HLT ADVANCEMENTS

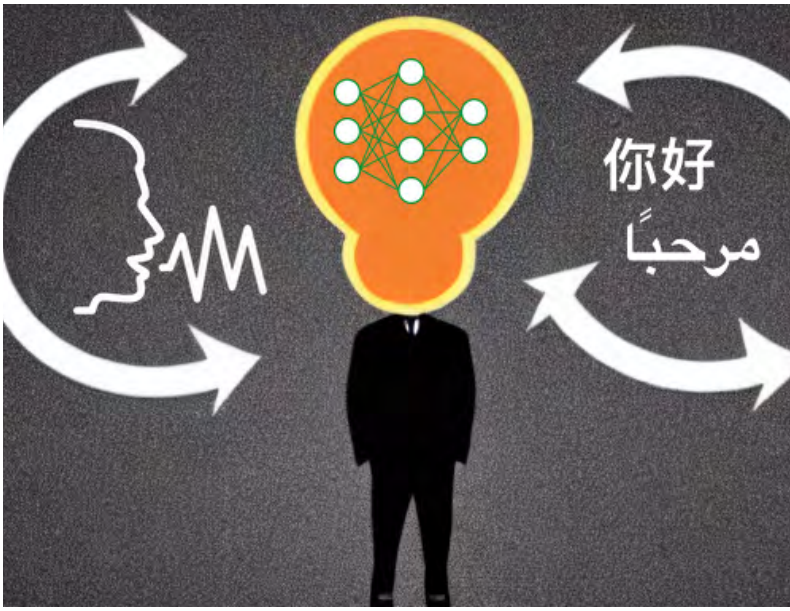
1) Successfully adapted and integrated a custom, license free, open-source Automatic Speech Recognition (ASR) system for the Network Integrated Training Environment (NITE) lab.

2) Showcased the ability for rapid response when operators reached back to AFRL for resources in Ukrainian ASR and Machine Translation (MT).

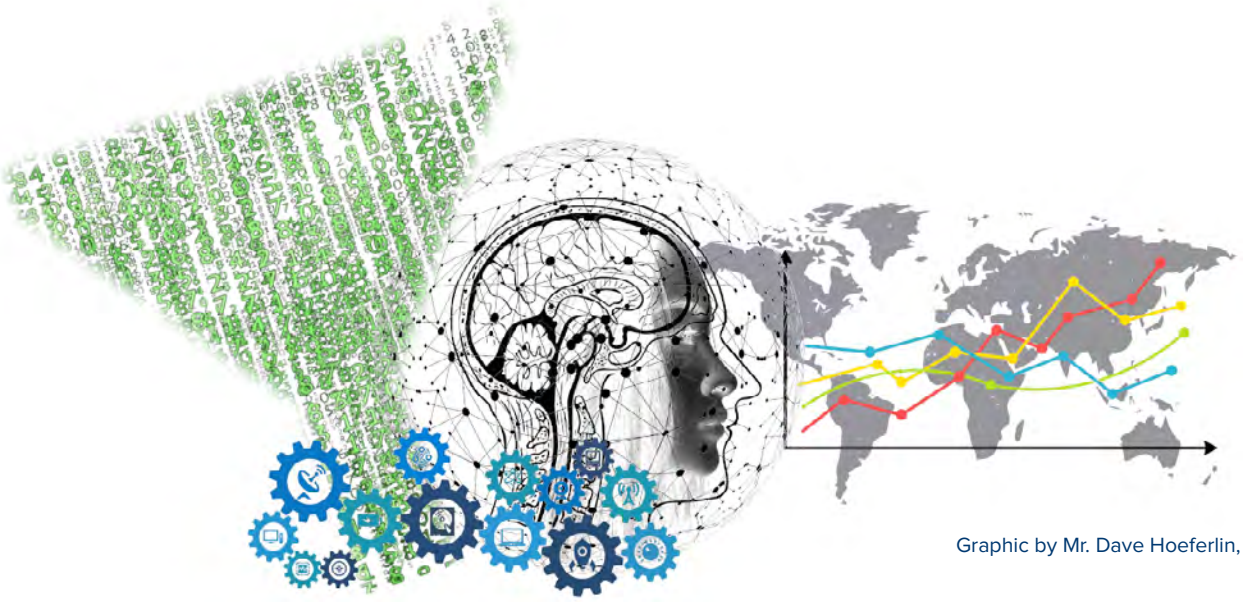
3) Invigorated previous efforts in Text To Speech (TTS) synthesis with the exploration of Neural TTS methods.

4) Explored multiple fronts of methods in Multimodal Learning, as an expansion to the current expertise of the HLT Group.

5) Leveraged technology received from other agencies, such as DARPA and IARPA, along with in-house developed capabilities, and has integrated and fielded this technology into a Cross Lingual Information Retrieval system.



Graphic by Stable Diffusion Image Generation and Mr. Jeremy Gwinnup, 711 HPW/RHWAI



Graphic by Mr. Dave Hoeferlin, 711 HPW/RHWAI

TASK 6: 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, with the support of advanced analytics. DWADE is split into two components, design and assessment. DWADE's design component seeks to develop a library of generalized data representations that can be adapted to specific work domains, like intelligence analysis. DWADE's assessment component is pioneering new methodologies to experimentally measure Human-Machine Teaming (HMT) effectiveness in applied systems. While both components are being executed in specific applied settings, DWADE's goal is to demonstrate generalizable principles and methodologies to be applied across the Air Force's mission.

DWADE's design component is currently centered on Project Hound. Hound is building support tools for intelligence analysts confronting data overload from ever-increasing sensor coverage and sensor diversity. Legacy tradecraft for geospatial intelligence exploitation does not scale with the growing availability of data, and is still attempting to better integrate multidisciplinary teams. 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 assessment component is developing methods to measure the impact decision support tool design can have on the human-machine team. This work is built upon the Ohio State University's Joint Activity Testing (JAT) methodology. JAT measures how HMT effectiveness degrades as challenge escalates by: identifying factors that make cognitive work difficult; generating scenarios that incorporate those factors; and running experts through those scenarios with the help of the decision support tool being tested. The resulting Joint Performance Graphs (JPG) detail how HMT performance drops off in the face of the various challenge factors. This data can provide valuable feedback to technology developers, helping them understand how to improve their tools to support the warfighters' cognitive work more effectively in the real world.

The DWADE team, in collaboration with the research arm of the National Geospatial Intelligence Agency (NGA-R) have recently finished their first study. This study measured the effect that recently developed decision support tools, designed to use machine learning analytics to help organize extremely large datasets, had on analysts' ability to answer intelligence requests quickly and efficiently.

Both components of DWADE are designed to work in concert. As both continue to mature, the effectiveness of tools created as part of the design component can be quantified by the methodologies created in the assessment component. ☆

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

DESIGN COMPONENT

Seeks to develop a library of generalized data representations that can be adapted to specific work domains, like intelligence analysis.

ASSESSMENT COMPONENT

Pioneering new methodologies to experimentally measure Human-Machine Teaming (HMT) effectiveness in applied systems.

TASK 7: JOINT ALL-DOMAIN INTEGRATED INTELLIGENCE, SURVEILLANCE, AND RECONNAISSANCE (ISR) (JADII) RESEARCH

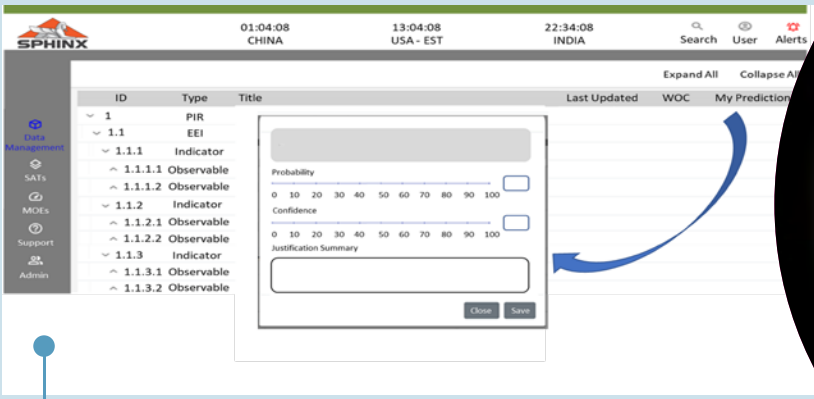
The 711th HPW/RHWAS Joint All-Domain Integrated ISR (JADII) Team recently began shifting its focus from individual elements of collection management to providing research and development (R&D) on the next generation of ISR orchestration and sense making. This has resulted in a New Start 6.3 effort called Legion. Previous groundwork was laid with investments in Kraken (Modeling and Simulation based ISR feasibility engine), Nautilus (Dynamic disposition of force and Tactics, Techniques and Procedures capability) and Sphinx (Collaborative Predictive Analytics) research. This has resulted in incredible customer engagement, external funding from Air Force and Intelligence Community partners, and functional capabilities that are being used by operators today. Legion represents a holistic framework, extending previous JADII projects to provide an end-to-end solution for the Sensing Grid. Our goal is to enable systematic identification of gaps in information, intelligent planning for deliberate and dynamic operations, and execution of missions with clearly defined closed-loop measures of performance. The Legion effort is directly supporting the United States Secretary of the Air Force (SECAF) Operational Imperatives 2 & 4, and is tied to the Air Combat Command/Intelligence Directorate's (ACC/A2's) functional requirements for the sensing grid.

Intel analysts and strategic planners involved with ISR operations are rapidly becoming overwhelmed with the large quantity of incoming information that must be collected, interpreted, and disseminated in near-real time. These data collections are essential in understanding enemy movement and resource allocation patterns to drive leadership's recommendations in both friendly and hostile environments.

To combat this issue, the JADII Team has developed a capability called Sphinx, a technology that facilitates the breakdown of complex problems, probing questions, and intelligence gaps to improve future prediction recommendations and situational awareness. Sphinx is led by Dr. Justin Nelson who has conducted several exploratory research studies to gather and collect insight into the tool's usability and capability, user-interface design, and decision-making efficacy. Within the last year, these research studies have provided invaluable information allowing the JADII Team to develop Sphinx in a manner that will best support intel analysts in the field.

Most recently, a scenario was developed and evaluated by 16 intel analysts at the 178th Intelligence Surveillance and Reconnaissance Group (ISRG) Ohio Air National Guard. The 178th intel analysts were excited a tool is being developed that will support their mission objectives. One analyst stated "I believe this tool will be extremely useful for strategic and tactical level operations." The commander of the 178th requested that our team return in the future to test and evaluate additional configurations of Sphinx and other capabilities from Legion. In addition, a small business innovation research (SBIR) phase II was awarded focusing on incorporating artificial intelligence and machine learning methodologies to bring critical information commonly overlooked to the forefront of the user. ☆

Dr. Jerred Holt, JADII Lead, 711HPW/RHWAS
Dr. Justin Nelson, Sphinx Lead, 711HPW/RHWAS



User Interface of Sphinx representing the workflow breakdown structure of Priority Intelligence Requirements (PIRs) into Essential Elements of Information (EEI), Indicators, and Observables to support future military predictions and recommendations.

Graphic and Screen Capture by 711 HPW/RHWAS JADII Team

COLLABORATIVE INTERFACES AND TEAMING

CORE RESEARCH AREA

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

TWO PROJECT AREAS

- **Human-autonomy Collaboration**
- **Distributed, Heterogeneous Teaming Solutions**

HUMAN-AUTONOMY COLLABORATION

COLLABORATIVE INTERFACES RESEARCH

The importance of autonomy for realizing USAF employment of multiple manned and unmanned vehicle (UV)-teamed sensor platforms in future warfighting is well recognized. These 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 to facilitate synchronized tactical operations. AFRL has already made significant advances in developing intelligent agents and ground station command and control interfaces that will enable human-autonomy teaming to manage multiple assets. Steps are now being taken to expand these capabilities beyond a single operator by designing systems to support teams of warfighters working together to accomplish shared mission goals. The approach utilizes a play-based delegation architecture, an intelligent Task Manager Interface, and a suite of interfaces to support the transfer of tactical control of assets - all designed 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 has examined 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. Our current effort aims to evaluate various team structures, interfaces, and system designs to best support and improve teaming between multiple human members, as well as between human-autonomy teams. This requires developing new capabilities and creating/enhancing display and control interfaces, including mechanisms to establish and modify working agreements between a human operator and an intelligent agent. Doing so will support effective teaming interaction in single and multiple IMPACT stations for multi-domain command and control of manned and unmanned assets. Part-task, lower-fidelity simulation, usability evaluations, and full scale experimentation with the high-fidelity IMPACT system are being used to assess candidate features, alternative interface designs, and decision aids. The results of iterative design and evaluation will feed technical efforts addressing 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

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

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

Screen Capture by 711 HPW/RHWC

The Sandbox Display is our main portal where the operator executes the mission – teaming with the human through a naturalistic dialogue with the autonomy via Intuitive Human Machine Interfaces (HMI).

Screen Captures by Dr. August Capiola

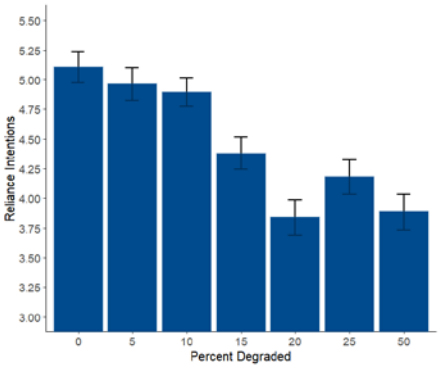
TRUST IN SWARMS

Robotic swarms comprise assets operating autonomously through local control laws. Swarm performance is difficult to assess, making appropriate levels of trust toward swarms critically important, especially when these systems behave unexpectedly or are subject to countermeasures. As part of the Center of Excellence for Trusted Human-machine Teaming—comprising 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 regarding relying on swarms. Our research showed that degradations to one facet of a swarm pervade into human beliefs about the entire system, which supports the notion of system-wide trust in human-swarm interactions. In 2022, we addressed several limitations from our initial data collections. We published data from a large online sample showing humans struggle to detect degraded assets in swarm foraging scenarios (Capiola et al., 2021). A follow-on data collection replicated and extended this work by offering a novel cluster analysis technique, which correlated highly with objective swarm degradation (Capiola, Johnson, et al., 2022). This suggests cluster analysis provides an interface affordance and viable decision-support tool to aid degradation detection in human-swarm interaction. Finally, we completed a data collection replicating and extending findings on the effects of asset degradation on trust toward swarms controlling for swarm performance, automation cues, and other features instantiated in our past work (Capiola, Hamdan, et al., 2022). Our results replicated and extended our earlier findings (while tightening up our experimental design) and have implications for quantifying human biases in human-autonomy interaction as swarm systems come online in AF-relevant contexts. ☆

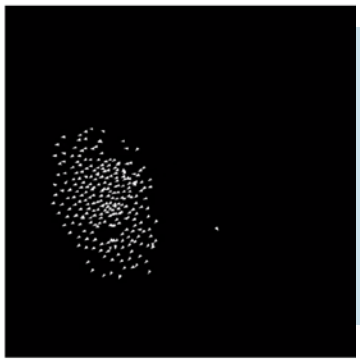
Dr. August Capiola, Research Psychologist, 711 HPW/RHWCT

Actual Degradation					10 Seconds After Degradation				End Degredation		
Percent of Assets Degraded	Time	Number of Clusters	Number of Outliers	Assets in Each Cluster	Time	Number of Clusters	Number of Outliers	Assets in Each Cluster	Number of Clusters	Number of Outliers	Assets in Each Cluster
0	0	NA	NA	NA	NA	NA	NA	NA	1	92	164
5	20	1	64	181	30	1	87	168	1	87	168
10	18	2	51	190_2	28	3	79	169_3_5	2	85	166_5
15	10	1	42	226	20	1	64	192.00	3	91	153_8_4
20	13	2	60	224_2	23	3	72	173_7_4	2	93	159_4
25	15	1	46	208	25	3	72	179_3_2	4	91	154_4_4_3
30	16	1	77	179	26	2	107	145_4	1	118	138
35	19	1	62	194	29	5	97	145_2_6_3_3	5	96	138_3_7_3_9
40	13	2	36	217_3	23	6	77	159-4-5-4-3-4	4	92	150-4-6-4
45	13	1	38	218	23	4	68	176-4-4-4	4	90	135-20-9-2
50	20	2	67	87_2	30	4	94	152-4-4-2	4	94	152-4-4-2
55	10	2	31	223_2	20	3	67	175-8-6	4	102	139-7-3-5
60	12	1	47	209	22	6	80	151-4-5-6-6-4	6	98	132-9-4-6-4-3
65	13	1	42	218	23	6	82	4-153-2-4-7-4	4	101	134-16-2-3
70	19	1	69	187	29	1	99	157.00	1	97	159
75	14	1	58	198	24	8	90	139_4_7_3_4_2_3_4	5	104	132_7_7_3_3
80	17	1	64	192	27	5	74	169_5_3_2_3	5	76	158_8_5_5_4
85	18	2	73	179_4	28	5	96	143_5_4_4_4	8	89	102_38_8_4_4_4_3_4
90	20	1	64	192	30	6	79	157_5_4_4_4_3	6	79	157_5_4_4_4_3
95	15	1	66	190	25	4	83	155_7_7_4	4	88	156_4_4_4
100	18	2	79	175_2	28	4	91	155_3_3_4	6	81	149_10_5_4_5_2

(Above) Depicts features of swarms experiencing varying degrees of degradation over a time, namely the relationship between the number of clusters at the end of each trial and the percent of assets that degraded.



(Left) Shows that there is a negative relationship between asset degradation and intentions to rely on swarms in a target foraging task, evidencing system-wide trust in human-swarm interaction.



From 0% - 100%, what percentage of the swarm do you perceive has degraded?

From 1 (Totally Uncertain) - 9 (Totally Certain), how confident are you in your estimate above?

(Above) Depicts a swarm experiencing 60% of their assets degrading over a 30-second trial. Participants responded when and the extent to which the degradation occurred, along with their confidence in said estimate.

Capiola, A. Hamdan, I., Fox, E. L., Lyons, J. B., Sycara, K., & Lewis, M. (2021). "Is something amiss?" Investigating individuals' competence in estimating swarm degradation. *Theoretical Issues in Ergonomics Science.* https://doi.org/10.1080/1463922X.2021.1983887

Capiola, A. Johnson, D., Hamdan, I., Lyons, J. B., & Fox, E. L. (2022, in preparation). *Detecting swarm degradation: Measuring human and machine performance.*

Capiola, A. Hamdan, I., Lyons, J. B., Alarcon, G. M., Lewis, M., & Sycara, K. (2022, revise and resubmit). *The effect of asset degradation on trust in swarms: A reexamination of system-wide trust in human-swarm interaction.*

Photo by Staff
Sgt. Andrew Lee

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

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 blackbox 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. This year we collected data on elegant failures in an experimental context. Preliminary results indicate elegant failures increased trustworthiness perceptions in the model over traditional ML models. ☆

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

WHILE YOU WERE LOCED:

REESTABLISHING SA WITH THE AID OF A MACHINE AGENT FOLLOWING A LOSS OF CONSCIOUSNESS EVENT

The USAF must harness advancements in machine learning and artificial intelligence technologies to quell anticipated force improvements of its adversaries. Competitive air defense can be achieved by the deployment of artifical intelligence/machine learning (AI/ML) enabled agents across a variety of AF missions to increase the number of friendly assets. Another important use case of AI/ML enabled agents is in the support of situation awareness (SA) maintenance in our pilots as they face increasingly complicated mission environments that require them to work with and monitor these agent teammates. In these environments, there is the potential for exogenous or endogenous disruptions to pilot SA, which could be offset by efficient presentation of rapid, streamlined displays of agent teammate intentions and actions. Such information may be most beneficial when a pilot is recovering from extreme physiologically adverse events (PE), such as gravity-induced loss of consciousness (G-LOC) or hypoxia. One of our aims at AFRL is to evaluate how a Pilot's Associate (PA) may accelerate the recovery of pilot SA following a PE.

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. In the past year, we have worked with industry partners to complete a cognitive task analysis with experienced military pilots that has revealed the task support needed in PEs and the preferred actions of a teammate, be they human or a machine. Our current survey research is aimed at developing more specific design aspects of the envisioned PA, such as modality for information delivery and what types of cognitive support provided by autonomy are desired by pilots. Findings will be implemented into a mission-relevant experimental study where we will observe the impact that 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 battlespace complexity, to moderate rapid and successful recovery of pilot SA. ☆

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

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

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

Photo by Master Sgt. Nicholas Priest

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

The work leverages previous research on human autonomy teaming (HAT) conducted at Clemson (McNeese), as well as at Georgia Tech and the Air Force Research Laboratory (Gorman and Myers, respectively), 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 plan a concentrated research effort that establishes a solid foundation to investigate the ability of trust to travel between teammates and between 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-method 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 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 organization-wide spread of distrust.

This proposed research will contribute fundamental knowledge to the general scientific community studying human-autonomy interaction, applied human-autonomy teaming domains, and current and future DoD initiatives. 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 often difficult to perform research that utilizes multiple HATs simultaneously. However, the teamwork experimental platforms available to the Principal Investigators, as well as their wealth of experience,

"This proposed research will contribute fundamental knowledge to the general scientific community studying human-autonomy interaction, applied human-autonomy teaming domains, and current and future DoD initiatives."

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

allow for multi-HAT constellation research to be conducted, serving as a landmark in the human-autonomy trust research domain. Third, we will consolidate the wealth of knowledge gained from multiple mixed-method experiments to provide an empirical tool specifically designed to measure the changes in one's trust of an autonomous teammate that are 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, Principal Investigator, 711 HPW/RHWM

Dr. Nathan McNeese, Principal Investigator, Clemson University (funded through AFOSR)

Dr. Jamie Gorman, Principal Investigator, Arizona State University (funded through AFOSR)

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 operators with different backgrounds and experiences to rapidly assemble into temporary teams, often with unfamiliar teammates. These distributed teams will have to coordinate across spatial distances and operational platforms to perform a mission and will afterwards disband. While the flexibility associated with temporary team formation may be necessary to meet the dynamic battlefield of the future, research shows that teams with temporary membership, sometimes referred to as ad hoc or fluid teams, often underperform compared to teams with stable memberships. This performance gap is due in part to teamwork factors that develop with team experience in stable teams, 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.

Our progress in this research area has followed two paths. First, an in-depth literature review of the factors and processes that influence teamwork in distributed teams has been undertaken and is nearly finalized. This review has resulted in a novel model of the contributors to distributed team success focused on individual-level, team-level, and team composition factors. The review also includes recommendations for selection, design, and training future fluid teams, as well as tools to assist the rapid formation of effective fluid teams.

Second, we have begun the development of a tool to assist fluid team member performance. To support this development, we have conducted interviews with several subjet matter experts (SMEs) to gather information and insight into the current and future distributed teaming needs of the Air Force. The tool will focus on presenting distributed team members useful information about the relevant experiences, knowledge, skills, and abilities of their teammates, and the mission-related capabilities they bring to the fight. Our goal for this tool is to support decision making, swift trust, and common ground formation to help rapidly close the performance gap between stable and fluid teams. A prototype of this tool is expected by the end of FY23. ☆

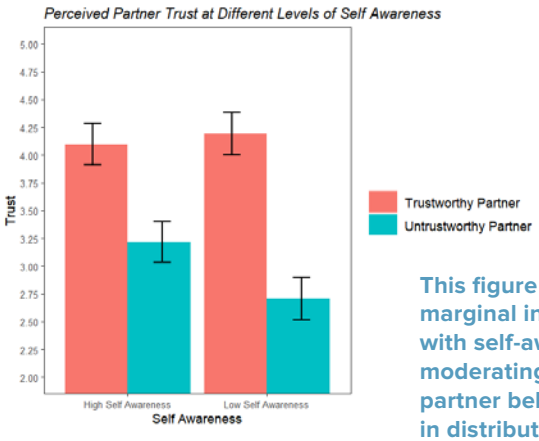
Dr. Michael Tolston, Research Psychologist, 711 HPW/RHWCT
Dr. Greg Funke, Senior Research Psychologist, 711 HPW/RHWCT
Dr. August Capiola, Research Psychologist, 711 HPW/RHWCT

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 22, we finalized data analysis from an online data collection leveraging a custom testbed to quantify the influence of differential risk (Capiola, Alarcon, et al., 2022). We also completed a data collection and are finalizing data analysis investigating the effect of self-awareness on swift trust (Capiola, Hamdan, et al., 2022), a novel antecedent identified in our task analysis of swift trust in joint all-domain command and control (JADC2) contexts. Analyses showed 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. Preliminary analysis shows a similar trend for partner purported self-awareness. Finally, we are gearing up to investigate the effect of partner familiarity—a heterogeneous feature in distributed military teams—and its interaction with partner behavior in distributed teams. Collectively, implications of this work concern the effects of antecedents to trust in distributed ad hoc teams as well as the role of familiarity in the trust process, the latter of which is a heterogeneous feature in JADC2 teams. ☆

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



Screen capture by Dr. August Capiola

This figure shows a marginal interaction with self-awareness moderating the effect of partner behavior on trust in distributed ad hoc teams (final analyses pending).

Capiola, A., Alarcon, G. M., Hamdan, I., Jessup, S. A., Ryan, T. J., Johnson, D., & Panson, D. (2022, under review). *What do I (and you) have to lose? Investigating the effects of risk asymmetry on the trust process in distributed dyads.*

Capiola, A., Hamdan, I., Harris, K., Alarcon, G. M., & Jessup, S. A. (2022, in preparation). *Investigating the moderating effects of self-awareness on the trust process.*

Screen capture by Dr. Kristen Liggett

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 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.

Since the creation of Merlin, significant progress has been made to include cyber effects in Joint All-Domain wargaming activities. The Merlin cyber module has been modified by RHWC human factors engineers to include an interactive user interface that supports the creation of a range of cyber effects to be played in a virtual wargame. In essence, players can use the graphical user interface to create a "quiver of cyber arrows" to have available over the course of the wargame, or before the wargame starts, allowing the player to conduct tradeoffs to maximize the number of effects available given available resources. The user interface was designed to support the generation of cyber effects by both cyber and non-cyber players, taking the mystery of "playing cyber" out of wargaming. Next steps in the project include the design of icons to call the cyber plays during the virtual wargame. ☆

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

TEAM PERFORMANCE METRICS

The Air Force relies on teams of Airmen for the success of critical missions with many demanding features, such as high tempo operations, long shift durations, and distributed team environments (e.g., joint all-domain command and control; JADC2). These demands can cause teams to underperform and jeopardize mission success. Therefore, developing metrics and methods to monitor, evaluate, and eventually augment teams in distributed settings is an important AF mission. At AFRL, we are accomplishing this by measuring a range of critical team processes, such as workload, situation awareness, and decision making, and have recently been focused on assessing joint attention as a foundation for estimating shared situational awareness. In this effort, we are using state of the art network-based methods to assess joint attention between individuals processing complex information. This approach may provide a powerful way to measure synchronous attention (e.g., to instructions, critical events, warnings, etc.) across multiple physiological and behavioral signals presented in distributed team settings. These methods may help to ensure that teams optimally process important messages and information, even when team members are operating in different domains or far away from one another (e.g., JADC2). ☆

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

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

> **...developing metrics and methods to monitor, evaluate, and eventually augment teams in distributed settings is an important AF mission."**
>
> **— Dr. Michael Tolston**
> Research Psychologist, 711 HPW/RHWCT

Graphic by Dr. Michael Tolston

Distributed team members are monitored for behavioral (e.g., communications and eye gaze data) and physiological (e.g., heart rate and brain activity) signals (top and middle panels). These data are then used to create networks that represent team interaction patterns (bottom panel). Finally, these networks are used to infer information processing properties and states of the distributed teams (e.g., joint attention).

REAL-TIME, OBJECTIVE MEASURES OF TEAM EFFICIENCY AND CALIBRATED TRUST

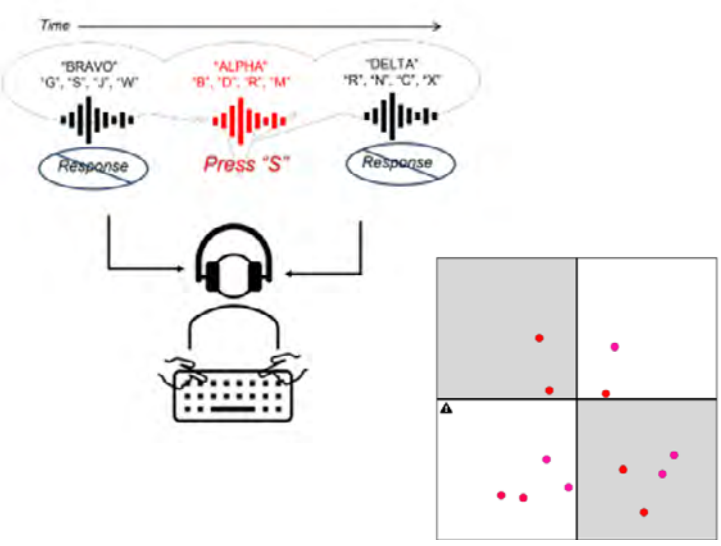
Trust affects performance and collaboration in risky, uncertain contexts. However, trust is predominantly measured with self-report scales or by proxy via risk-taking behavior after an interchange has occurred, both of which have limitations. To address these limitations, Drs. Fox and Capiola derived a novel, unobtrusive, and real-time metric of team trust and efficiency in complex, multisensory environments. To investigate the utility of this metric, the researchers first generated trust manipulations and synthetic agent performance, both of which were leveraged in a full laboratory experiment. Fox and Capiola found quantitative data supporting which trust manipulations are deemed relevant and influential in ad hoc teams (Capiola et al., 2022). In parallel, they generated agent behavior from pilot human-subjects data to manipulate agent performance (Fox et al., 2022). From there, the researchers conducted an in-person experiment investigating the interaction between trustworthiness and performance on trust calibration and team efficiency in ad hoc distributed teams, and data analysis is ongoing. Drs. Fox and Capiola also co-mentored a 2022 Repperger Summer Intern to begin work on designing an interface that leverages multiple modalities to dynamically communicate team efficiency in a way that demands the fewest competing attentional resources. Implications of this work concern transparent interfaces aimed to facilitate effective communication in ad hoc teaming scenarios. This work demonstrates cross-branch collaboration between the Multisensory Perception and Communication and Collaborative Interfaces and Teaming CRAs to investigate AF-relevant solutions for efficient collaboration in distributed ad hoc teams. ☆

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

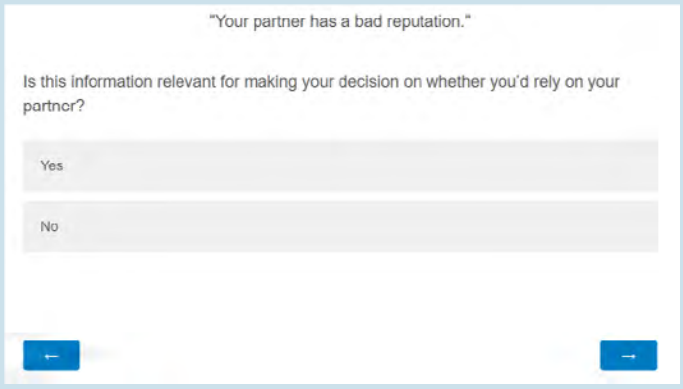
Capiola, A., Fox, E. L., Stephenson, A., & Hamdan, I. (2022, revise and resubmit). *"Is this even relevant?" Investigating the relevance of antecedents to trust in ad hoc dyads.*

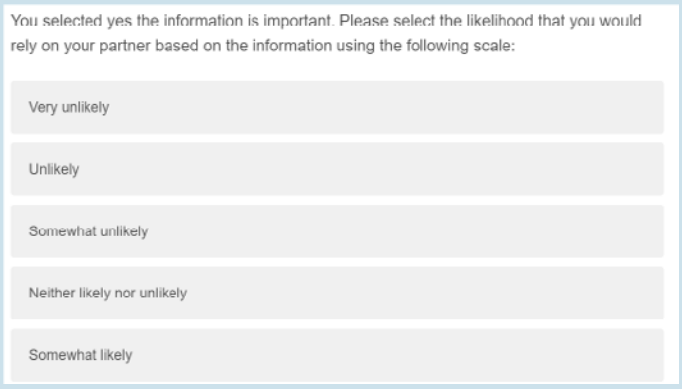
Fox, E. L., Bowers, G., Capiola, A., & Stephenson, A. (2022, revise and resubmit). *Designing an artificial teammate after human performance and interruption behavior in a multiple object tracking and communications task.*

(Below) Demonstrates the process by which researchers quantified the a) relevance and b) effect of antecedents on swift trust in ad hoc teams.



Screen Captures by Dr. August Capiola





(Above) Demonstrates the process by which researchers quantified the a) relevance and b) effect of antecedents on swift trust in ad hoc teams.

MULTISENSORY PERCEPTION AND COMMUNICATION

CORE RESEARCH AREA

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

TWO PROJECT AREAS

- **Exploiting Perceptual Abilities**
- **Enhancing Operational Communications**

Effective communication of information is critical for the success of every operator, in any domain, and in every Air Force and Space Force mission. These missions routinely take place in settings that are perceptually, cognitively, and environmentally complex. Operators are challenged by multiple sources of stimulation, intense task demands, and situation uncertainty. Each of these and other factors pose unique challenges for effective portrayal of complex information. The Multisensory Perception and Communication Core Research Area (MPC CRA) is focused on identifying and exploiting the underlying sensory and cognitive mechanisms mediating perception and communication in order to inform the development of multimodal interfaces and advanced communication technologies.

As FY23 unfolds the MPC CRA will begin to focus our research more exclusively towards challenges associated with Joint All-Domain Command and Control (JADC2) and Human Representation for Digital Engineering (HRDE). As part of this refocusing, several of the existing efforts under both the Exploiting Perceptual Abilities (EPA) and Enhancing Operational Communications (EOC) lines will ramp down as new efforts are stood up to address these important operational concepts.

New efforts planned for the FY23 include design and development of intelligent multimodal common operation pictures. This concept pulls requirements directly from the ABMS Tactical C2 Battlespace Awareness user interface concept where the key desire is to “display, analyze, and communicate” from within a common user interface. Efforts from JADC2 echo issues seen in the Base Defense and Dismounted Special Operations domain where information flowing through situation awareness tools does not stay in sync with information through traditional, natural language-based channels like voice and chat. Our efforts will examine this problem from the perspective of Human-Human (H-H) communication, where the common operating picture and language-based channels function as two modalities within a single, more holistic communication medium. This will inform how to embed ad hoc communication tools within traditional Common Operating pictures, and investigate features of tools like hypermedia tagging to provide cross linkages to information across the modalities.

Continued on next page

THE TWO LINES OF EFFORTS

Exploiting Perceptual Abilities (EPA)

Enhancing Operational Communications (EOC)

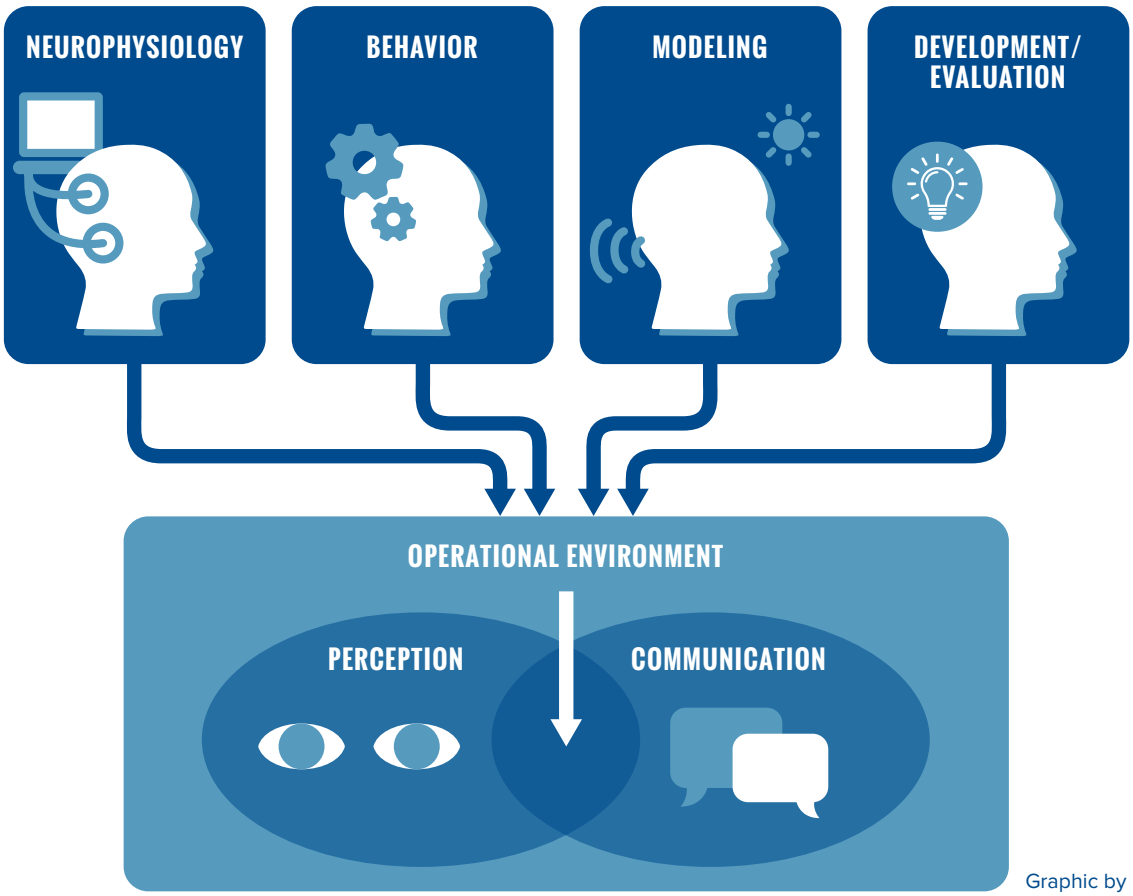
Another new area is designing and developing agile communication interfaces. This project focuses specifically on overcoming communication challenges with degraded and disrupted channels. We will pull in requirements from the Advanced Battle Management System (AMBS Communications Broker Needs statement, efforts will focus on using artificial intelligence (AI) and machine learning (ML)-powered natural language processing techniques to mediate communication over degraded channels. An example might be automatically switching a conversation from full motion video to voice chat in response to channel degradation without requiring a user to manually change channels or production methods. Guidance for this work stems directly from the human-human communication literature on multimodal grounding and media richness coming from our current communication resiliency task. ☆

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

Operators are challenged by multiple sources of stimulation, intense task demands, and situation uncertainty."

— Dr. Griffin Romigh
Multisensory Perception and Communication CRA 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 inundated with, and must effectively integrate, multiple, concurrent information streams. This is particularly true in military operational environments, where stimulus and task complexity are especially high and time constraints are severe. Operators have finite attentional resources to filter and process this information, so 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 actions of their teammates to maintain situational awareness and quickly intervene in response to unexpected events. The Managing Multisensory Information Throughput line of effort uses neurophysiological and behavioral measures to examine the strengths (and limitations) of human perception and cognition, and to define principles allowing proper integration and distribution of multisensory data in operational environments. We seek to design displays that tailor to individual operators, considering capabilities and proclivities of the individual, to ensure that information is processed in the most efficient way, and overall decision making and performance is optimized. To best do this, we examine how sensory information processes interact, starting from basic neurological and psychophysical principles, to 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; 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; 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; identifying neurophysiological markers and behavioral correlates of multisensory perception, attention, and decision processes, and distribution; the 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 multitasking strategies, inform interface adaptation, and supply pertinent feedback to users and teammates so they more effectively work together. ☆

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

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

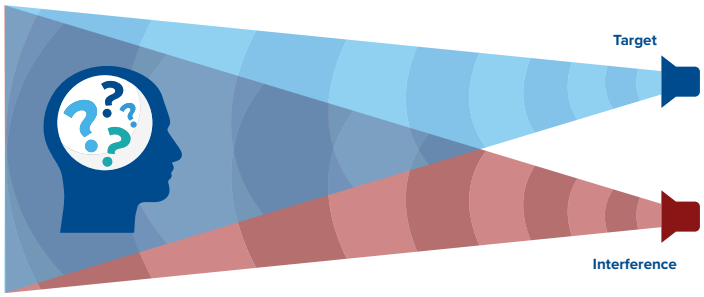
Graphic by Mr. Will Graver

INFLUENCING PERCEPTION

Humans have a remarkable ability to parse complex visual and auditory environments, forming perceptual objects out of the scene to extract relevant information. However, this process is imperfect and often incomplete, potentially leading to situations where critical information is missed. Understanding the abilities, as well as the constraints, of human perception and cognition informs the development of advanced capabilities that will enable the US Air Force 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 best to 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 increasing use in 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. Ultimately, 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



Graphic by Ms. Shania Horner

Photo by Tech. Sgt. Marelise Wood

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, in particular, are acoustically rich and complex, and may play a major role in the success, or failure, 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 may have a deleterious effect on advanced technologies such as automatic speech recognition to function effectively. Conversely, the acoustic environment may be exploited by operators, whether 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). Thus, when simulating operational environments for training, interface development, and teleoperation, it is critical to provide a veridical representation of the acoustic environment. In effort to, among other requirements, '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. 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 or attitude or air conditioning fan speed). To enhance realism and perceptual accuracy, communication signals will be processed such 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

Graphic by Mr. Will Graver

IMMERSIVE VIEWPOINT SHARING

Future Air Force operations will continue to grow in complexity, requiring coordinated activities across distributed agents that support missions 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 may 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 understand 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 in control of 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 ownship 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 and 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

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 a conduit for efficient information flow in human-machine teams. Research conducted under the Communication Resiliency line of effort 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 line of effort involves the development of a number of laboratory-based collaborative tasks that capture the dynamics of real operational communication environments, use of experimental paradigms to simulate conversational repairs and grounding in voice user interfaces, the development of spoken and natural language processing models capable of creating conversational alignment and recovery, and the evaluation of 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 humans use to circumvent communication challenges to inform the development of natural communication interfaces that will inevitably encounter similar communication errors specifically within complex and time-critical human-machine interactions. ☆

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

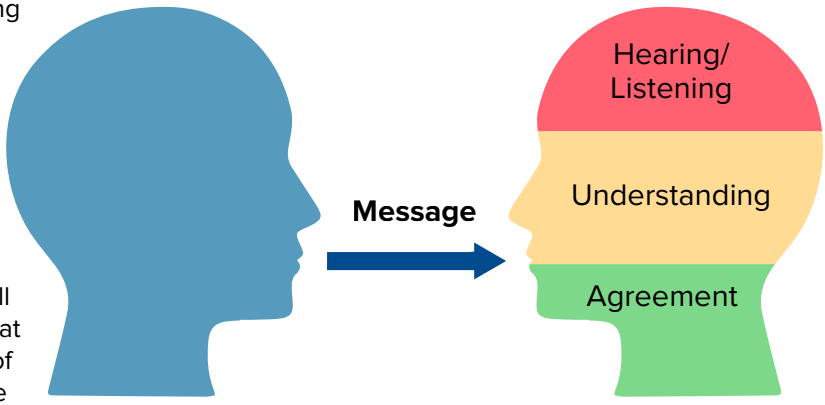
RESEARCH IS FOCUSED ON:

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

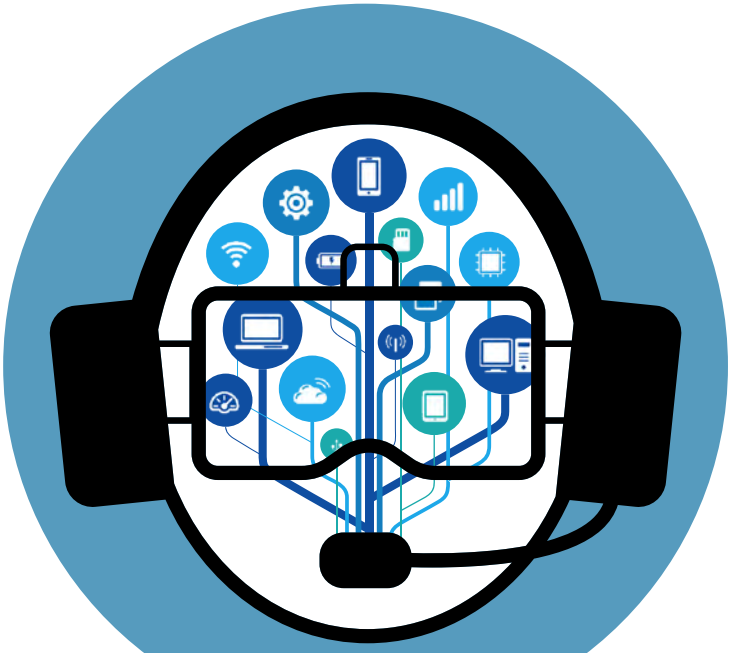


Graphic by Ms. Shania Horner

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 line of effort 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 as a result of a failure in the system's ability to effectively analyze an acoustic scene to segregate *target* from *masker*. Research conducted under this line of effort 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 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



Graphic by Ms. Shania Horner

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. Every mission relies on the effective communication between teammates and our operators serve in cognitively demanding roles where it will be increasingly critical to have communication interfaces that overcome, rather than contribute to, informational overload and the loss of situation awareness due to attentional tunneling. Our research aims to combine advancements in our understanding of human perception with recent developments in artificial intelligence (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 interaction with the communication channels themselves. This will allow operators to dynamically winnow non-critical communication traffic and receive the most relevant information at the times, and 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. ☆

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

CONTEXT-AWARE COMMUNICATION INTERFACES

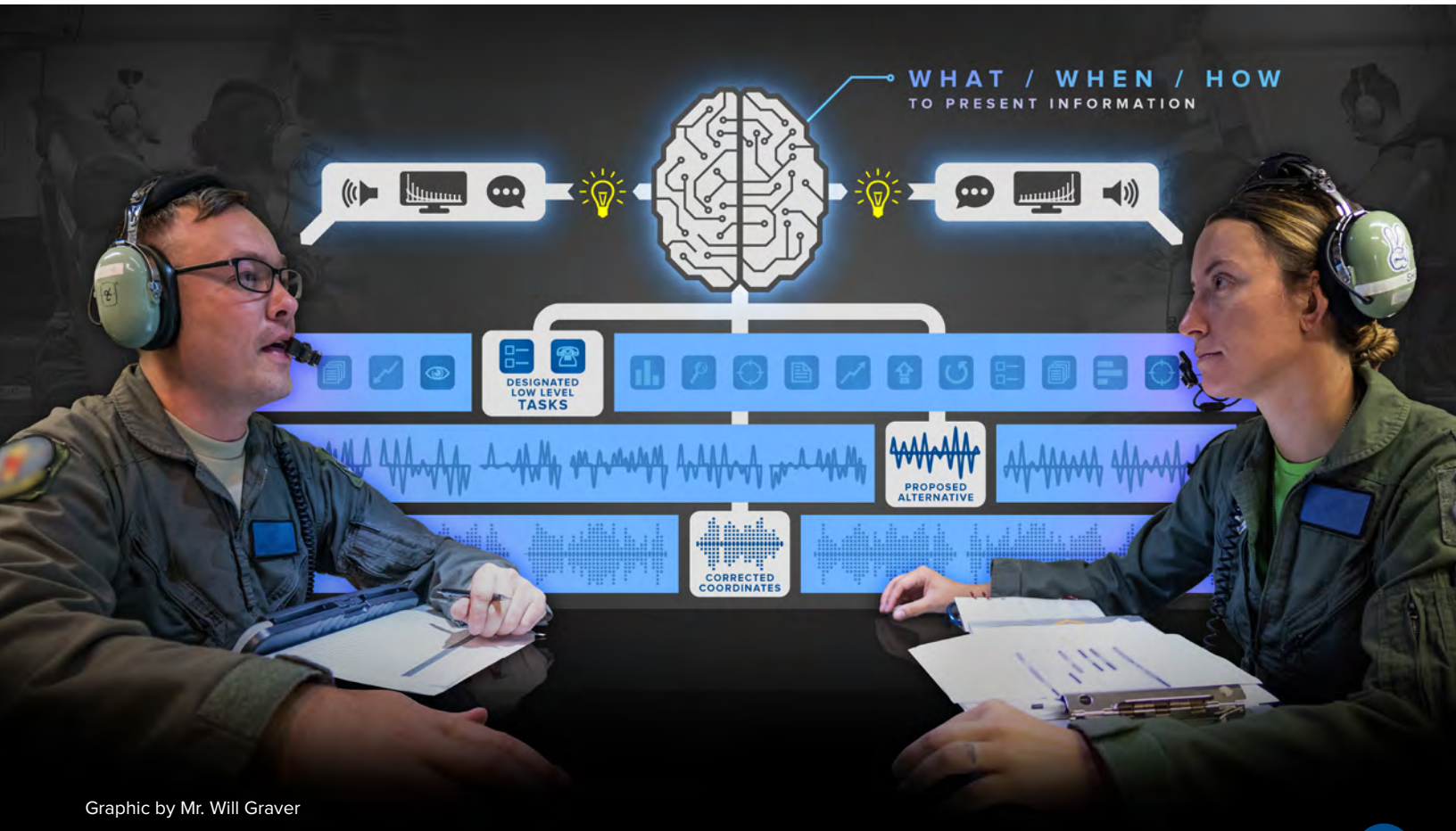
As the warfighting mission becomes more complex, operators are required to manage more information from disparate sources, attend to a greater number of concurrent tasks, and effectively collaborate with distributed team members across domains and mission sets. These military missions are likely to have time-critical components that rely on the expeditious presentation of up-to-date information to inform ongoing tasks. Thus, 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 line of effort 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 not only to automate low-level tasks, but may also 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 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, and involves 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 helps to 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

RHW FY22 SUCCESSES

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

Transformational Capabilities Office (TCO) Engagement

- Skyborg
- Rocket Cargo
- Fight Tonight
- Mayhem
- Resolute Sentry
- High-speed Vertical Takeoff and Landing (HSVTOL)

Low Altitude Sensing Helmet (LASH) Lysander Transitioned to AFLCMC Big Safari Program

Joint All-Domain Integrated Intelligence, Surveillance & Reconnaissance (JADII) ISR Feasibility and Optimization Engine "Kraken" Receives Approval for On-boarding to Air Force's Platform 1

Content Data Standards and Tools Transitioned to Air Combat Command Distributed Training Center

The **GRILL® Created DIS Game Engine Plugins** (Unreal, Unity, Cesium) That Greatly Improve Interoperability of Game-based Military Simulation with More Traditional DoD M&S Sim Environments

Completed a New IEEE/SISO Standard Called **Compressed Distributed Interactive Simulation (C-DIS)**

Delivered Cognitive Realism Improvements within **Advanced Framework for Simulation, Integration and Modeling (AFSIM)** Agents to AFSIM User Community

The Mission Readiness App Transitioned to Fatigue Optimized Cognition Under Stress (FOCUS) 6.3 program in Airman Biosciences Division with target customers Air Mobility Command and Air Force Special Operations Command

Predictive Analytics for Learning (PAL) **Transitioned to Air Education and Training Command Linguist Next, Defense Language Institute**

Adaptive Teamwork with Layered Airman-Machine-Interfaces & Systems (ATLAS) and **Skyborg Human Systems Integration (HSI) Team** Completes First Integration with Test Pilot School X-62 VISTA Platform Linking the Flight Test Engineer with Autonomy

Networked Integrated Tactical Environment (NITE) Joint Fires Transitioning to JSOC

Intelligent Multi-UxV Planning with Adaptive Collaborative Control Technologies (IMPACT)/Multi-Environmental Domain Unmanned Systems Application (MEDUSA) Integration

EST. 2001

To view more issues, visit:
www.afresearchlab.com/FightsOn

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-2022-5313

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