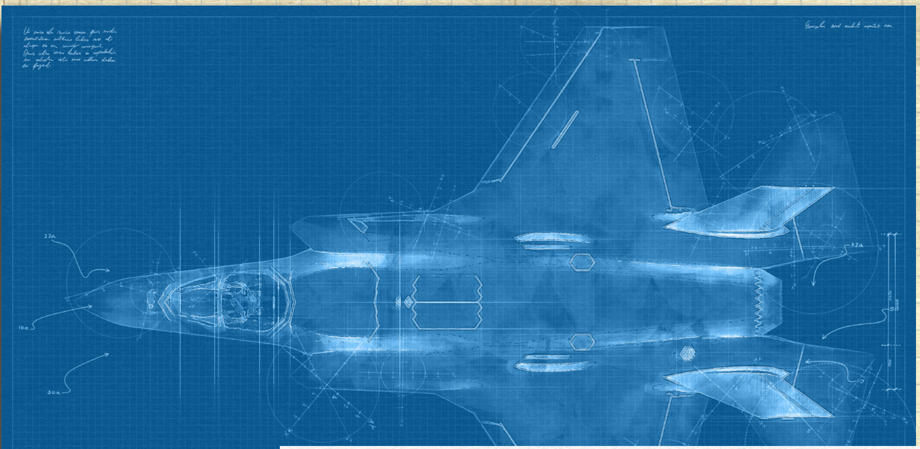
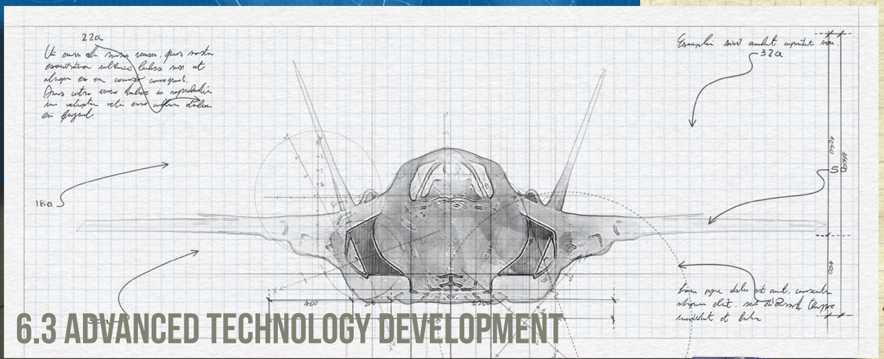
CALL SIGNS

MARCH 2022
Call Signs Volume 10, Issue 1

USN AEP
SOCIETY



6.1 BASIC RESEARCH



6.3 ADVANCED TECHNOLOGY DEVELOPMENT

6.7 OPERATIONAL SYSTEMS DEVELOPMENT

The RDT&E Spectrum
from concept to completion

How research and funding go hand-in-hand

RDT&E OVERVIEW

A quick primer on how the US Government funds research and development, and how the work gets done.

PAGE 4

HORIZON OF PREDICTION

Can a machine predict something as complex as human behavior? We take a look at ground breaking research into the science of post-traumatic stress discorder

PAGE 8

CRACKING THE CODE

Traumatic brain injuries account for a significant amount of disabilities in US Servicemembers, but tracking exposures has been impossible... until now.

PAGE 14

CONTENT

12 DIET PERFORMANCE

How athletes and US military members heal from traumatic injuries may be determined, in part, by what they eat. This research explores how low-carb, high-fat diets may encourage cellular regeneration and speed healing.

16 LASER SAFETY

What the Naval Aeromedical Research Unit is doing to protect aviators from laser strikes.

20 AUTONOMOUS CASUALTY EVACUATION

Preparing for tomorrow's battlefield conditions requires innovative thinking, and lots of testing. We explore the role that autonomous vehicles may play in casualty evacuation through hands on field tests with the United States Marine Corps.

30 HYPOXIA TRAINING

Learning to detect and mitigate the effects of low oxygen (hypoxia) when flying at high altitudes is critical to pilot safety. Read how one group of AEPs has developed a new training device to help aviators do just that.

33 AVIATION SAFETY

Aviation mishap rates have been on a steady decline for decades, but lingering accident rates continue. What we attribute to improved safety and what is actually the root cause may surprise you.

37 MEET AN AEP

What do you want to do when you grow up? For Sarah Beadle, AEP #164, the answer was simple.

39 USNAEPS AWARDS

Read about the accomplishments of AEPs around the community and learn more about the amazing work being done across the US Navy.

FROM THE SPECIALTY LEADER

AEPs Past, Present and Future:

I hope you had the chance to take some time off during the recent holidays, to relax and recharge. In my last note, I shared with you the unique role that AEPs play in ensuring our Naval Forces are manned, trained, equipped and ready to win any mission, beat any adversary.

While our Aeromedical Research and Acquisition expertise is key to maintaining this role, our experiences ensuring that the Naval Aviation Enterprise (NAE) can operate in complex, high-hazard domains for extended periods without serious accidents or failures allows us to support Navy Medicine's High Reliability Organization (HRO) mission. In that vein, I am honored to share with you that a team of ten AEPs were awarded the 2021 RDML Lewis E. Angelo Symposium's "High Reliability Organization Poster Showcase" award (non-clinical category). The work summarized in this poster represents more than a decade's worth of innovation and collaboration across the Navy/Marine Corps, Army and Air Force to deliver an Unmanned Air Vehicle Operator Selection capability. These findings, coordinated with the NAE, the Naval Medical Enterprise and the Bureau of Naval Personnel, were key to the recent establishment of the Navy's FIRST EVER Warrant Officer Unmanned Air Vehicle Operator community.

This is just one example of the many different and impactful contributions our community has made over the past year. AEPs continue to provide Operational Medicine solutions, aligned to HRO principles. We seek new ways to deliver Human System Integration - enabled, Acquisition – driven capabilities. We do this while remaining committed to driving innovative and groundbreaking research.

Collectively, AEPs authored over a dozen peer-reviewed articles, were invited to serve as guest editors to several scientific journals, presented and chaired sessions at numerous conferences and, gathered a range of military and industry organization awards recognizing the quality and impact of their scientific and technical work.

Beyond these achievements, you'll find many of us in key leadership roles across the Medical and Aviation enterprises, leveraging our expertise to ensure our Naval Forces are ready and lethal.

Moving into my second year as Specialty Leader, I could not be prouder of all we have done together this past year and am looking forward to sharing even greater successes this year.

CAPT Joseph Cohn, AEP #113

FROM THE PRESIDENT

The United States Aerospace Experimental Psychology Society (USNAEPS) Executive Committee (EXCOM) is pleased to present the latest issue of Call Signs. This past December, I was honored to be selected to serve as USNAEPS President. On behalf of USNAEPS, I would like to express our sincere appreciation to CDR Brent Olde for continuing the tradition of excellence set by all previous USNAEPS Presidents.

In addition to CDR Olde's departure from the EXCOM, we are grateful to the outgoing members whose countless hours of hard work continued to keep the society strong and made Call Signs, AEP Annual Meetings, the USNAEPS Awards Program, and our recruiting efforts possible. We are pleased to announce that that LCDR Mike Natali will continue serving our society, moving from Treasurer to Vice-President. Our Historian CAPT(Ret) Frank Petho and Member-at-Large, CAPT(Ret) Mike Lilienthal will also continue to serve USNAEPS. We are also excited to announce new additions to the EXCOM team. LT Nick Armendariz will assume responsibilities as Treasurer, LT Claire Modica has filled the vacant position of Secretary, and LT Sarah Sherwood has taken over as Editor, replacing LCDR(sel) Todd Seech. I would be remiss if I did not take the opportunity to express our collective gratitude to outgoing Vice-President LT Eric Vorm. Since 2012, Eric has served our society in a variety of roles. He has enthusiastically tackled tasks such as updating www.navyaep.com, crafting and executing our social media strategy, and managing the layout and graphic design for Call Sings. In addition to Vice-President, three new EXCOM members will fill Eric's shoes. LT Adam Braly is serving as our new webmaster, LT Alexandra Kaplan has taken on the challenge of Social Media Coordinator, and LT Aditya Prasad has officially assumed the responsibilities as Recruiting Coordinator.

In addition to new EXCOM members since the last edition of Call Signs, our community is thrilled to welcome LT Nick Armendariz (AEP # 163), LT Sarah Beadle (AEP # 164), LT Alexandra Kaplan (AEP # 165), and lateral transfer from the Aviation Community, LT Jennifer Knapp. LT Knapp took the Oath of Office as a Medical Service Corps Officer on March 4th and she will be awarded AEP number 166 upon completion of the Aeromedical Officer Course later this year.

This issue of Call Signs presents a small sample of research led by AEPs that ranges from Basic Research (6.1) to Operational System Development (6.7) as well as impactful research conducted outside the realm of Research, Development, Test and Evaluation (RDT&E) activities. Our Specialty Leader, CAPT Joseph Cohn, starts this edition off with an overview of the Research Acquisition Lifecycle. From there we feature two 6.1 efforts by LT Eric Vorm and LT Claire Modica that respectively investigate the use of machine learning to predict Post Traumatic Stress Disorder and novel nutritional approaches to improve human performance and recovery. LT Modica also presents Applied Research (6.2) that characterizes blast exposure to help mitigate traumatic brain injuries. Next, LT Sarah Sherwood's team describes an Advanced Technology Development (6.3) effort that resulted in a capability to mitigate laser threats to aviators. Moving along the continuum, LT Vorm presents an excellent example of Technology Demonstration and Evaluation (6.5) in his assessment of an Autonomous Casualty Evacuation capability for the Marine Corps. The last example of AEP work in RDT&E is a project that LCDR Brennan Cox and I initiated as a Management and Support (6.6) effort. This work successfully transitioned to Operational System Development (6.7) and delivered an Aircrew Task for use in each of the Navy's eight Normobaric Hypoxia Trainers. CDR Jeff Grubb departs the realm of RDT&E to describe impactful research conducted at the Naval Safety Command which is poised to reshape how the Navy understands, educates, and prevents aviation mishaps. Finally, we meet one of our new AEPs, LT Sarah Beadle, who recently completed aviation training and earned her "wings of gold."

On behalf of the USNAEPS EXCOM and all of our contributing authors, I hope you enjoy and widely share this issue of Call Signs and THANK YOU for your continued support of the United States Aerospace Experimental Psychology Society.

LCDR Lee Sciarini, AEP #141

FROM THE BENCH TO THE BATTLEFIELD

An overview of the Research, Development, Testing and Evaluation (RDT&E) Framework

By: Captain Joseph Cohn, PhD
Aerospace Experimental Psychology Specialty Leader

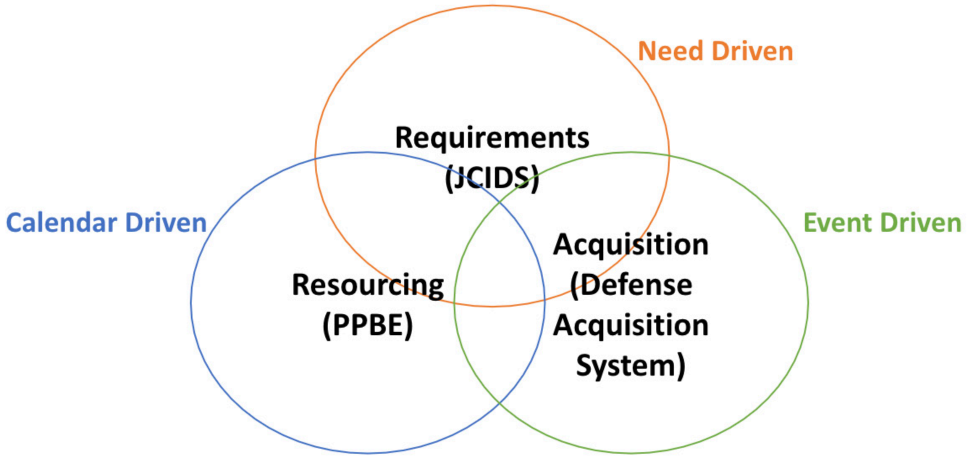
As uniformed Health Care Scientists within the U.S. Navy's Medical Service Corps, Aerospace Experimental Psychologists (AEPs) play key roles across the Research, Acquisition and Sustainment (RAS) lifecycle, shepherding innovative ideas from the bench to the battlefield. In order to bridge the gaps that naturally occur at the seams between Research, Acquisition and Sustainment, it is imperative that we understand the three pillars that underlie the RAS lifecycle: Requirements, Resourcing and Acquisition. In this overview, we will provide a short discussion on each of these pillars and how they impact the overarching RAS Lifecycle.

The three pillars are often considered as intersecting processes that are both interdependent and independent of each other. The pillars are interdependent in that they each inform the other. For instance, the Acquisition pillar requires focus, provided by the Requirements pillar. The Requirements pillar articulates needs to be addressed, through the Acquisition pillar. Both Acquisition and Resource pillars require Resources to execute their respective processes. At the same time, the pillars maintain a level of independence from each other. Resourcing follows a series of calendar-driven processes, collectively known as Planning, Programming, Budgeting and Execution (PPBE). Requirements follows a need-driven process, known as the Joint Capabilities Integration and Development System (JCIDS). Lastly, Acquisition follows an event-driven process, moving from one development stage to the next only after planned milestones for a given stage are determined to be complete by a designated Milestone Decision Authority (MDA). Let's look at each pillar in turn, starting with requirements.

While Requirements may derive from many different sources, Capability Based Assessments (CBA) are typically the starting point, within the JCIDS process, for building a complete requirement. CBAs generally involve a group of scientists, acquisition professionals, subject matter experts and users, convened for the purposes of exploring an emergent challenge, tasked with: quantifying the challenge in terms of specific capabilities that are needed; characterizing at the high-level parameters of those capabilities; and quantifying gaps to close in order to deliver those capabilities. Other requirements sources may include Wargames or Joint Capability Technology Demonstrations (JCTD).

Regardless of starting point, these activities represent a typical approach to defining a long-term "future state." These types of requirement development efforts typically assume a long lead time (more than two years); anticipate significant research investments; and lead to "traditional" requirements documents, like Initial Capabilities Documents and so forth. Recognizing that the operational environment is dynamic, JCIDs supports other, quicker ways to derive and act on requirements. These include Joint Urgent Operational Needs (JUONs) and Joint Emergent Operational Needs (JEONs). Both reflect requirements that must be addressed in less than two years, leading to an accelerated acquisition process, and differing primarily in terms of the anticipated impact to mission readiness if the corresponding capability is not delivered.

The Acquisition process, recently updated during Calendar Year 2020 (https://aaf.dau.edu/aaf/policies/) follows an event-driven approach to understanding and addressing a given requirement. The process typically allows for initial Research investments, which help refine an understanding of the requirement. These investments are often at the Basic, Applied or Advanced Technology Development levels, depending on the maturity of the Science and Technology domains needed to address the requirement. These investments provide the input necessary to negotiate with the MDA-specific criteria to meet in order to proceed with developing a materiel solution to address the gap, and to determine where in the Acquisition pro-



The process of researching and developing technology requires considerable planning. The RDT&E framework provides structure to guide research efforts towards success.

cess the next stage should begin.

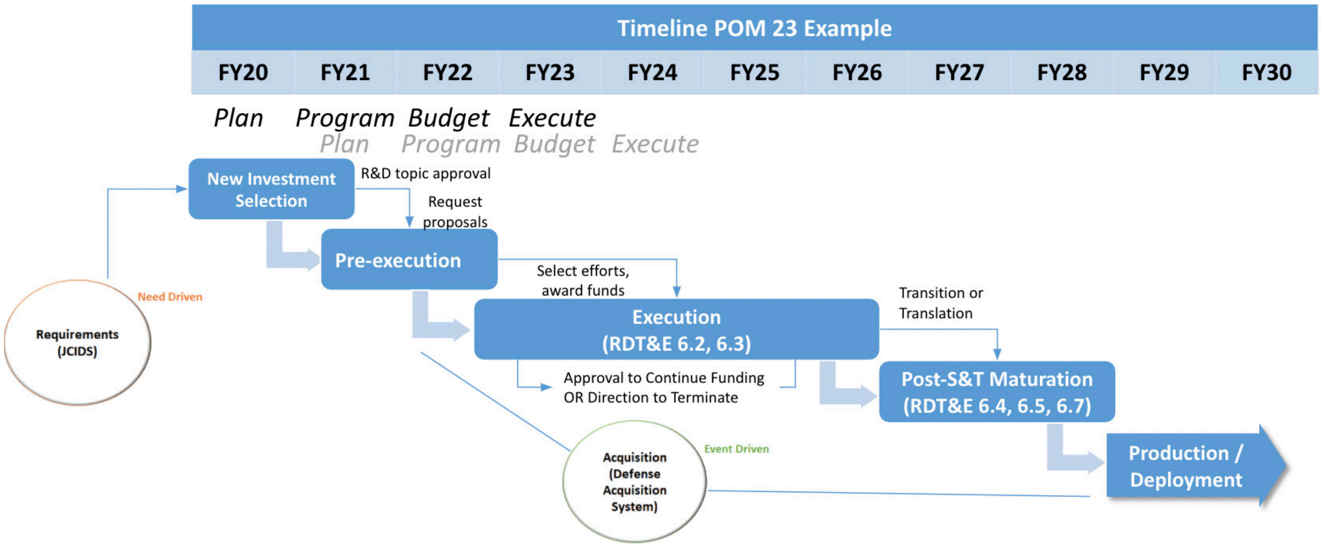
This "Material Development Decision (MDD)" marks the formal entry of the effort into the first step of the Acquisition process: the Materiel Solution Analysis. In some instances - where the technology is more mature and the risk is already manageable–the beginning point could be a later stage in the process. Regardless of entry point, once started, funding appropriate to the Acquisition stage must be identified (typically ranging from Advanced Technology Development to Advanced Component Development to System Development funding types) and a series of performance criteria agreed to, in order to progress to each successive stage, or Milestone, via a Milestone Decision by the MDA.

The Acquisition process allows for a level of flexibility and tailoring appropriate to a specific need and capability, taking into consideration the expertise of the teams supporting the effort, the level of maturity of the capability's underlying technology and the Warfighting population for whom the capability is intended. The recent updates to Acquisition guidance, which have led to the "Adaptive Acquisition Framework" (https://aaf.dau.edu/) provide an even greater level of tailoring to more effectively accommodate specific capabilities, needs and technologies.

The final pillar, Resources, is tightly connected to a series of calendar driven activities – Planning, Programming, Budgeting and Execution. Together, these activities support the allocation and application of the resources necessary to deliver new capabilities that address specific requirements. The planning activity lays the foundation for the Department of Defense, and Service and Defense Agency specific, strategic focus areas through a review of relevant guidance. The resultant guidance informs high-level areas in which the Services and Defense Agencies will either continue to invest, or from which they should divest. During the programming activity, Services and Defense Agencies develop their resource requirements, known as the Program Objective Memorandum (POM – based on the planning activity outcomes—for a five-year period called the Future Year Defense Program; FYDP). These resource requirements feed into the budgeting activity, with each Service preparing a Budget Estimate Submission (BES) for the first year of the FYDP. The last activity in this process, execution, is where the funds that have been programmed and budgeted are applied to efforts that will address guidance developed during the planning phase.

In any given year, all four of these Resource processes are active. To illustrate how this pillar evolves over time, consider the phasing of the four PPBE activities when developing the Fiscal Year (FY) 2023 POM. Here, the planning activity begins in FY 20, so that funds are available for execution in FY 23. Moving into FY 21, we continue with the programming of funds to address the plans developed in FY 21. At the same time, we begin planning for the FY 24 POM. As we continue forward, we can see how each successive FY requires accounting for multiple future years' effort.

Pulling together these three pillars provides a sense of how we, as uniformed scientists, can identify challenges facing our Warfighters and plan to deliver



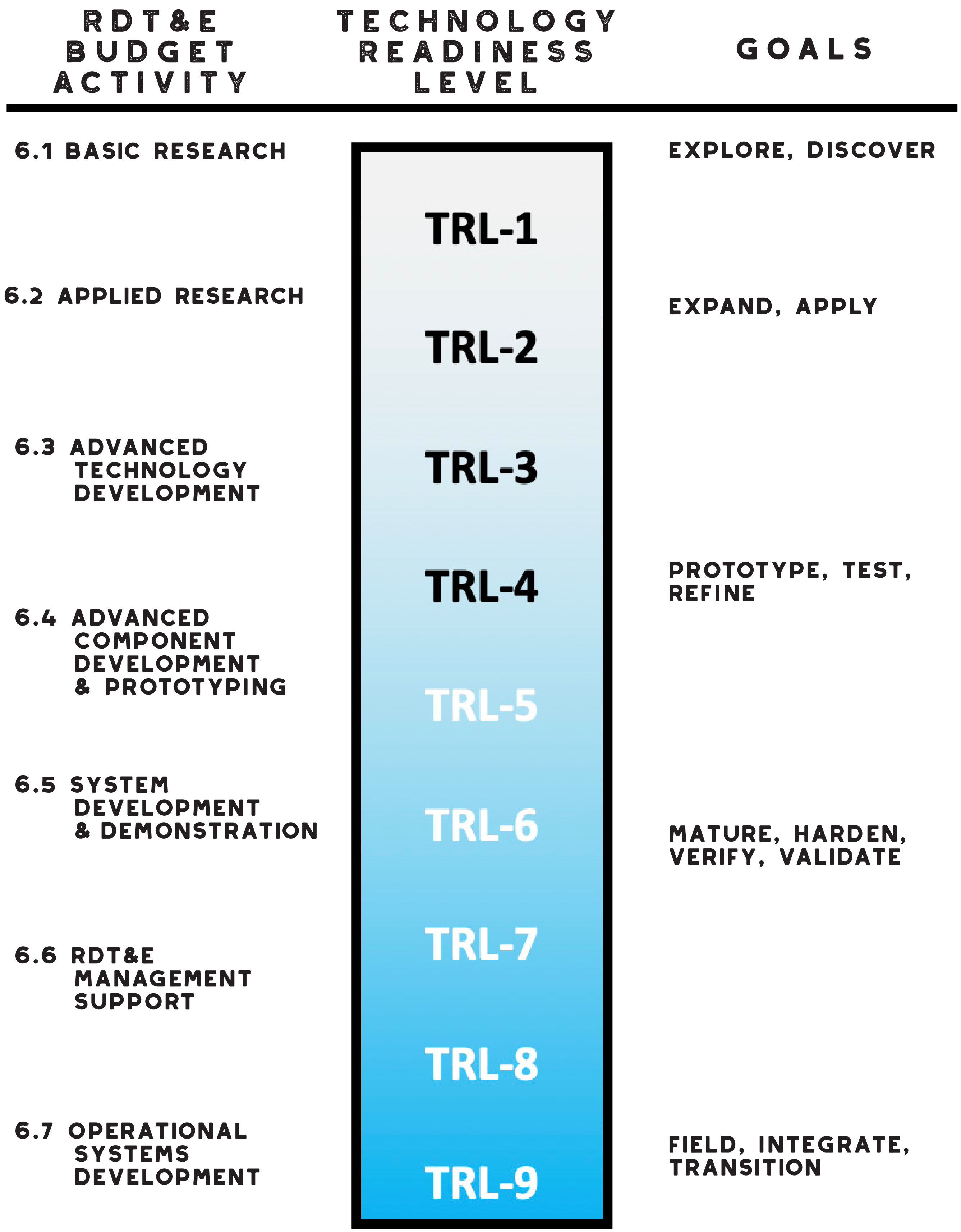
Timeline POM 23 Example

FY20	FY21	FY22	FY23	FY24	FY25	FY26
Plan	*Program*	*Budget*	*Execute*			
	Plan	*Program*	*Budget*	*Execute*		
		Plan	*Program*	*Budget*	*Execute*	
			Plan	*Program*	*Budget*	*Execute*

to them solutions, over the long term. Using the POM 23 example, in FY 20 we start by identifying a challenge. These challenges may align to an existing requirement or, if not, could then provide the basis for coordinating input to generate a new requirement. With that information, we can begin to plan to address that requirement – to include understanding the necessary research. With that understanding, we can work with leadership in the following FY (FY 21) to program funds for research to address that requirement. Moving out one more FY, we have one more opportunity to confirm the budget necessary to solve the challenge before moving into FY 23, the actual year in which funds begin to be executed to develop that solution.

In summary, moving our research from the bench to the battlefield requires more than simply being experts in our respective scientific fields. It demands that we be prepared to effectively plan across the three RAS Pillars. It means that we must understand the requirements that clarify the challenges our research is addressing; it mandates that we play a key role in advocating for the resources we need; and it requires that we take a flexible approach to tailoring how we acquire the needed capability. In short, understanding how to integrate the three lifecycle pillars is critical to transforming our research from the bench to the battlefield.

Complex projects like the F-35 Joint Strike Fighter (above) require tremendous science and technology efforts to bring them into operational existence. Basic science, such as exploring the behavior of new materials at high heat and supersonic speeds (6.1), is then translated into new high-performance components (6.2) which are ultimately incorporated into a fully-working system with new capabilities (6.7). Scientists and engineers receive funding according to the type of work they are doing and its ultimate purpose in order to conduct the research and develop new components and systems. **US Navy Aerospace Experimental Psychologists** are involved in every aspect of these efforts, from conducting basic research on things like human performance in high-G environments, to managing teams of contractors doing work on behalf of the US Government, to coordinating the successful transition of new technology to the fleet. The RDT&E spectrum is the framework that guides all of these efforts.



Developing new or improved capabilities to the point where they are appropriate for operational use is the goal of research, development, testing and engineering (RDT&E). Each RDT&E appropriation is divided up into seven budget activities, which are designated 6.1 - 6.7. The type of work being done and the immediate goals of the research and development efforts determines what type of money (budget activity) can be used to accomplish that work. Scientists and engineers such as US Navy Aerospace Experimental Psychologists write proposals and compete for funding across the entire RDT&E spectrum. The above graphic provides a quick reference to understand how budget activities align with technology readiness levels. The graphic also broadly describes the purpose of the work corresponding to each level.

HORIZONS OF PREDICTION

Can we leverage the power of Machine Learning to predict complex psychobehavioral phenomena?

By: LT E.S. Vorm, PhD
Deputy Director, Laboratory for Autonomous Systems Research

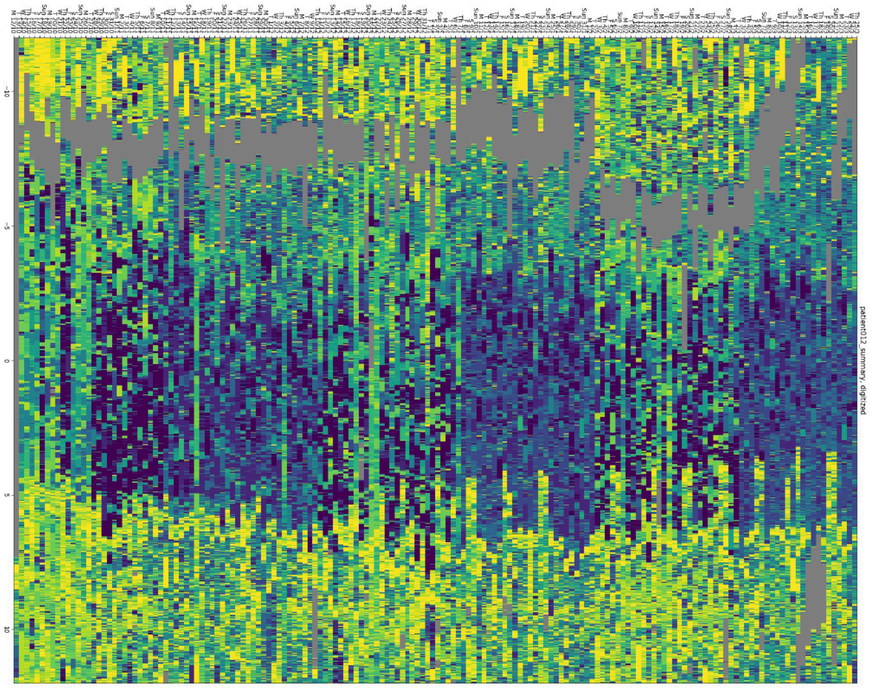
Popular depictions of military research and development are abundant. Unfortunately, however, visions of military scientists in white lab coats crafting futuristic weapons in batcave-like laboratories have an unintended side effect: the expectation that scientific breakthroughs happen frequently and immediately. The reality, unfortunately, is that most research is more mundane and time consuming. The building blocks of technologies like radar-deflecting stealth panels or laser-guided munitions come in the form of small but significant scientific baby steps, documented in technical reports, conference proceedings, and scientific journals. Vannevar Bush, first director of the Office of Scientific Research and Development and the person who ran the military's research and development during WWII, once famously said that basic research "creates the fund from which the practical applications of knowledge must be drawn" (1, Ch. 3). The US Government's strategy to invest in basic research represents its understanding of the fundamental relationship between scientific knowledge and its offspring practical applications. This article highlights one example of basic research (6.1 on the RDT&E funding spectrum) in the realm of military medicine. To set the stage and introduce the topic, let's begin with a brief scenario.

A Marine helicopter is performing routine training when suddenly one of the engines experiences a catastrophic failure. Despite the heroic efforts of the pilots, the helicopter crashes. Four of the six crew members on board, including both pilots, die in the crash. The remaining two crew members are both injured and are taken to the hospital where they both spend several weeks recovering from their injuries.

A year later, one of the surviving crew members has returned to her job and is experiencing no serious symptoms of prolonged stress after the accident. She has moved on from the traumatic event, and shows signs of continuing growth. She has strengthened her relationships with her loved ones and family, and has renewed her interests in hobbies such as playing ultimate frisbee and long-distance bicycling. Because of the positive interactions she experienced with her caregivers while in the hospital, she has begun taking college courses and is preparing to apply to a nursing program. She has a renewed sense of appreciation for life, and seeks to be mindful and thankful in her daily life, activities, and relationships.

The other crew member, however, has not fared as well. He has been experiencing serious adverse symptoms of prolonged stress since shortly after being released from the hospital. He experiences frequent nightmares, and is easily startled and reacts uncontrollably to being surprised to the point of becoming enraged. He angers easily, and is increasingly irritable. Depression, anxiety, and a prolonged sense of dread make it increasingly difficult for him to function in social circumstances. Because of frequent outbursts and unstable behavior at work, he was removed from flight status and his security clearance was frozen pending a medical review. Eventually his symptoms become so severe that he is admitted to an inpatient facility to manage his ever-growing inability to cope with daily life after he tells co-workers that it would have been better if he had died in the crash and that he wishes he was dead on a regular basis. He is eventually deemed no longer fit for active-duty service and is ultimately medically separated from the Marine Corps.

What factors determine who will recover from trauma and who will experience prolonged psychological stress? Are these factors learned? Can they be trained? Or are these factors genetic? Questions such as these have interested human beings for as long as we have recorded human history. People as far back as the ancient Greeks noticed and commented on how some soldiers are able to experience traumatic adversity with resilience and carry on their lives after war without issue, while others seem to be permanently scarred and altered by their experiences. Despite many decades of modern-day research in genetics and the psychology of stress, however, little is known or understood about the complex phenomenon known as post-traumatic stress disorder (PTSD).



Our algorithm learns four separate scores from the data gathered with our wearable device. Each score is associated with a depressive characteristic: diurnal activity, anxiety, psychomotor retardation, and circadian rhythm (sleep). The image on the left is all the data from a single patient across a period of approximately six months. Patients wear the device while in treatment. Their activities and physiological data are recorded and used to create our predictive algorithm.

Traumatic events such as combat exposure, near-death experiences and sexual assault affect every person in some way. As the Psychiatrist Viktor Frankl once put it, “an abnormal reaction to an abnormal situation is normal behavior” (2, p. 20). Contrary to popular notions of PTSD, however, while many people develop short term effects (e.g., sleep loss, mild anxiety symptoms) following a trauma, most people recover within a short period of time, and relatively few people develop long-term PTSD. The trick to preventing serious PTSD is to identify early those who are more vulnerable and susceptible before chaotic expressions of PTSD are allowed to fully develop. But as we saw in the vignette earlier, predicting who is more or less likely to develop long term psychological effects of trauma is very difficult. As a result, organizations such as the US Department of Defense incur extremely high costs in terms of manning and medical treatments related to PTSD in service members.

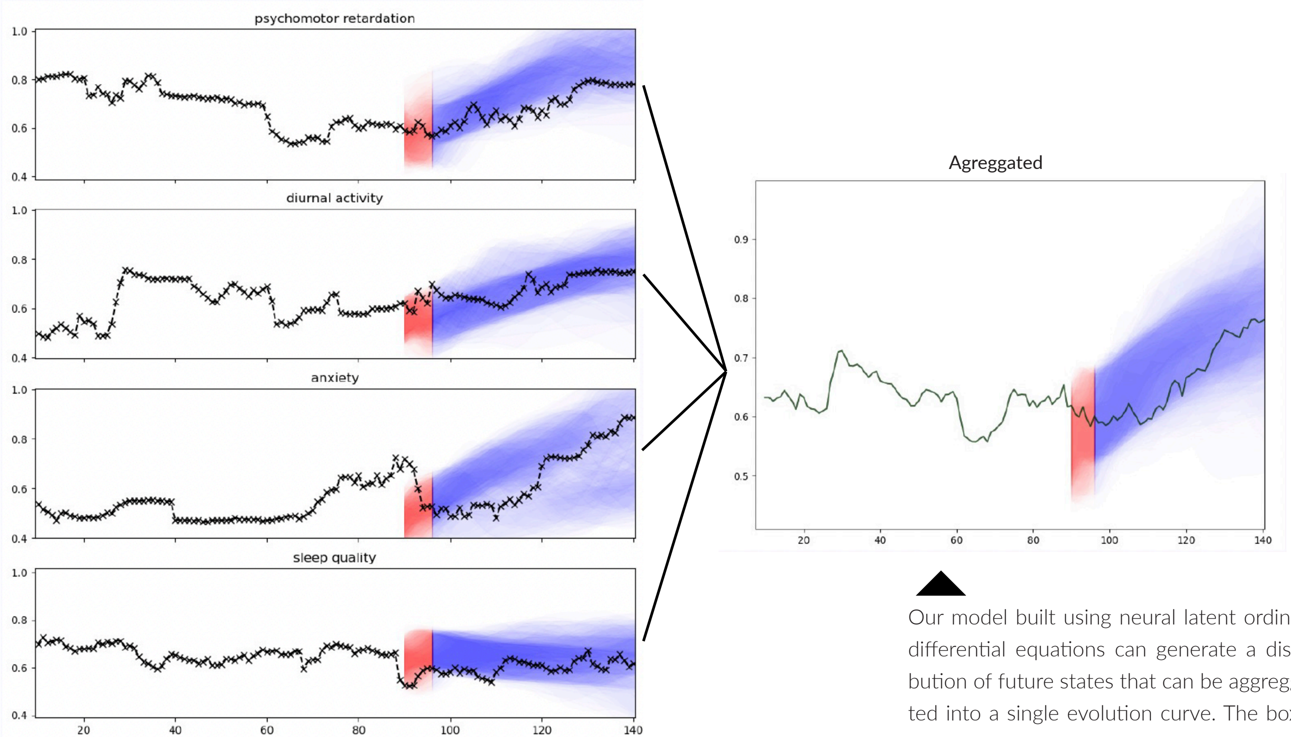
Estimates of the number of service members experiencing PTSD are alarming—between 13.5% and 30% of recently deploying troops have tested positive for PTSD and required significant medical treatment (3, 4). This number totals over 500,000 troops over the past 13 years of conflicts in the Middle East (5). The total costs of this is in the multiple billions of dollars, but even greater are the costs in terms of retention of qualified military personnel and medical readiness.

One of the most common aims in PTSD research, therefore, has been to create models of patient trajectories that could serve as an early warning for patients who are likely to need more help recovery from a traumatic event. Patient models could indicate whether patients are improving or whether they are declining and moving towards full PTSD in ways that current medical practices do not afford. Clinicians could use these models to inform their treatment decisions, which would then improve patient outcomes. And today there is reason to be optimistic that model-based approaches to predicting PTSD might be possible thanks to recent advancements in artificial intelligence.

AI is everywhere today—in our cities, in our homes, and in some cases even in our bodies (6, 7). Due to the recent explosion of computer processing speed and power, machine learning (ML) approaches to data science have yielded exciting new opportunities to learn more about ourselves and the world around us than ever before. ML’s advantage comes from its ability to perceive relationships between variables at extremely high dimensions. This means that ML can learn patterns and associations between variables at levels far beyond human comprehension, and in a fraction of the time it would take for humans to perform those same calculations. ML has been used to successfully predict astonishing things with remarkable accuracy, from discovering new drugs (8) to predicting who might live or die in the intensive care ward of a hospital (9). Most recently, my team and I began to wonder whether ML could predict something as complex as PTSD, and whether the future Department of Defense could use predictive analytics to provide early warnings for personnel who are most at risk at developing severe PTSD. Inspired by this question and armed with some very bright minds and powerful computers, we set out in early 2020 to answer this basic question.

Through a cooperative agreement between the U.S. Naval Research Laboratory and the Office of Naval Research Global, teams of scientists and mathematicians from France and the U.S. teamed up to explore how machine learning could be used to predict PTSD. In order to utilize machine learning, we first must have enough data to train the system. To get this data, our first step was to utilize an existing bespoke wearable device to facilitate data collection. The wearable that we chose had been developed in partnership with the Office of Naval Research Global, and has already been approved for human participants experimentation through the l’agence nationale de sécurité du médicament et des produits de santé (ANSM), which is the French equivalent of the U.S. Food and Drug Administration. This wearable device combines a photoplethysmography (PPG) meter, with an actimeter and electrodermal activity (EDA) meter. Each of these represents the state-of-the-art in wearable technologies and facilitate things like detailed sleep analysis, evaluations of stress, overall physical activity, blood pressure, blood flow, and oxygen saturation. The device is meant to be worn 24 hours a day, and continuously collects data for months at a time.

With this device we collected six months’ worth of patient data through a cooperation with local treatment centers in



Our model built using neural latent ordinary differential equations can generate a distribution of future states that can be aggregated into a single evolution curve. The boxes on the left represent the four variables we are interested in, collected from our wearable device. The box on the right is the agreggated score. To the right of the red line is projected data based on the data to the left of the red line. In this example, the model was trained on the first 90 days of data, and then asked to predict future states. The data is robust and reliable out to several weeks in the future. Using a system such as this, clinicians could possibly detect the onset of severe conditions or relapse early enough to intervene--something that is impossible to do today.

France, as well as participation with the French Army. We collected data from 200 patients who had been diagnosed with major depressive disorder, 200 patients who had been diagnosed with PTSD, and approximately 2,600 healthy patients.

One challenge in modeling PTSD is that it has an enormous range of expressions, meaning that two people with identical diagnoses may look very different in terms of the symptoms they are experiencing. The range of symptoms and tendency of those symptoms to be subjectively derived via self-reported instruments (i.e., surveys filled out by patients and interpreted by clinicians) further complicates the use of sophisticated mathematical models such as those used by machine learning. Rather than attempting to model and predict PTSD directly, therefore, we determined that we would first attempt to model and predict another malady that frequently occurs in coordination with PTSD—major depressive disorder (MDD). It is estimated that between 50-70% of patients diagnosed with PTSD are also diagnosed or diagnosable with MDD (10). In addition to this significant overlap, major depressive disorder can be more readily identified and diagnosed through the use of physiological data, such as the types we were collecting. One additional benefit to attempting to predict MDD for this project was that patients in our study were also evaluated by licensed clinicians using the Montgomery-Asberg depression rating scale (MADRS). MADRS is a questionnaire that patients complete at various intervals while in treatment, and is used by clinicians to document and quantify the severity of symptoms. Each patient in our sample received six MADRS evaluations, one per month. With this MADRS data, we constructed patient trajectories, representing ground truth for how each patient fared during our data collection period. This ground truth served as the baseline against which our team could evaluate the accuracy of our machine learning algorithm.

With these trajectories constructed and our data collected, we trained our semi-supervised neural network and began to explore how it learned what features (e.g., heart rate variability, sleep disruption, daily step count) provide the best predictive power, and how accurately we could predict patient outcomes with the data we collected.

To do this we fed our neural network the first 60 days' worth of patient data, and then asked it to predict patient trajectories for the remaining portion of our six-month window. We then compared these machine-generated trajectories with the ground truth we constructed from MADRS data. The results were very robust.

As you can see from the diagrams Agree, the trajectories correlate very highly with MADRS data, which is a strong indication that this automated approach to detecting complex psychological suffering is entirely feasible. Most notable of our findings is that we were able to make accurate predictions of patient trajectories (e.g., who is likely to relapse, who is likely to recover) with only 60 days of data. This means that a clinician, armed with data from a wearable and an algorithm like ours, could potentially intervene weeks or months before symptoms become severe. This is of immense practical value if you consider that the current standard of practice for diagnosing and quantifying major depressive disorder (i.e., MADRS) can only describe how a patient is feeling at any given moment, but cannot accurately predict what a patient will do in the near or far-term future. Even the most skilled and seasoned clinician is likely unable to make accurate predictions about patient outcomes, especially when dealing with complex pathologies such as major depression or PTSD. As you can see in the figure below, a clinician treating this patient at Time 50 may believe this patient is getting significantly better, when in reality they will soon experience a significant relapse and a return to clinically significant symptoms shortly after this data point. It is worth repeating that our algorithm accurately predicted this trajectory, despite only having the first 60 days of data. Only through the added computational power of machine learning can clinicians hope to gain an edge in forecasting future patient states.

We have demonstrated the feasibility of an approach by utilizing technical and scientific expertise. Out of this small study may grow a technology that could one day greatly reduce the long-term effects of trauma, but there are many other studies necessary before the U.S. Department of Defense might be willing to embrace this approach (or one like it) at full scale. Such is the nature of basic 6.1 research. Through a small investment in time and resources, we have grown our knowledge and explored something new. The next step in the RDT&E journey (6.2) would be to expand this research and focus it further.

If you would like to read a more in-depth account of this research, you can read our article published in the International Journal Human-Intelligent Systems Integration:

Fompeyrine, D.A., Vorm, E.S., Ricka, N. et al. Enhancing human-machine teaming for medical prognosis through neural ordinary differential equations (NODEs). Hum.-Intell. Syst. Integr. (2021). https://doi.org/10.1007/s42454-021-00037-z

REFERENCES

1- Bush, V. Science The Endless Frontier. A Report to the President by Vannevar Bush, Director of the Office of Scientific Research and Development, July 1945. United States Government Printing Office, Washington: 1945.

2. Frankl, V. E. (1984). Man's search for meaning: An introduction to logotherapy. New York: Simon & Schuster.

3- Eber S, Barth S, Kang H, et al. The National Health Study for a New Generation of United States Veterans: methods for a large-scale study on the health of recent veterans. Mil Med. 2013;178:966–969.

4- Tanielian T, Jaycox LH, editors. Invisible Wounds of War: Psychological and Cognitive Injuries, Their Consequences, and Services to Assist Recovery. Santa Monica, California: RAND Corporation; 2008.

5- Thompson M. Unlocking the secrets of PTSD. Time. 2015; 185:40–43

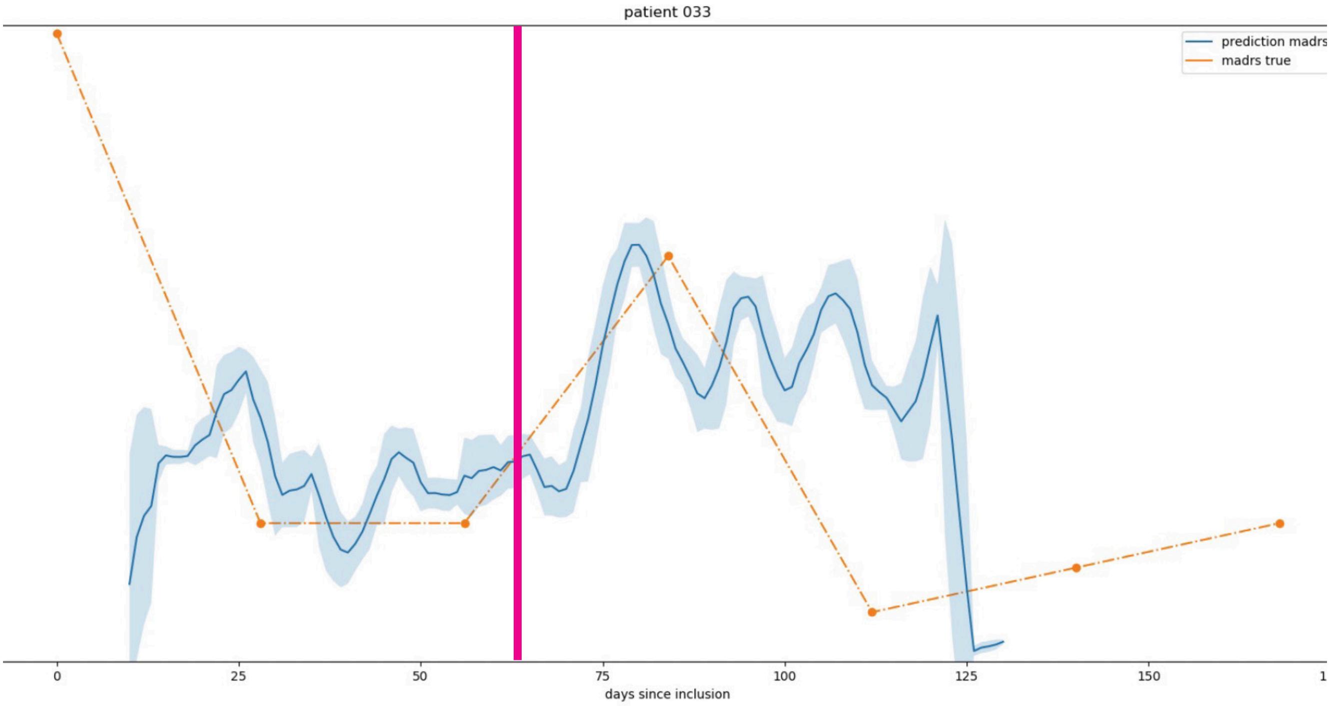
6- Reardon, S. AI-controlled brain implants for mood disorders tested in people. Nature 551, 549–550 (2017). https://doi.org/10.1038/nature.2017.23031

7- A. N. Balaji and L.-S. Peh, "AI-on-skin: Enabling On-body AI Inference for Wearable Artificial Skin Interfaces," Ext Abstr 2021 CHI Conference on Human Factors of Comput Systems, pp. 1–7, 2021, doi: 10.1145/3411763.3451689.

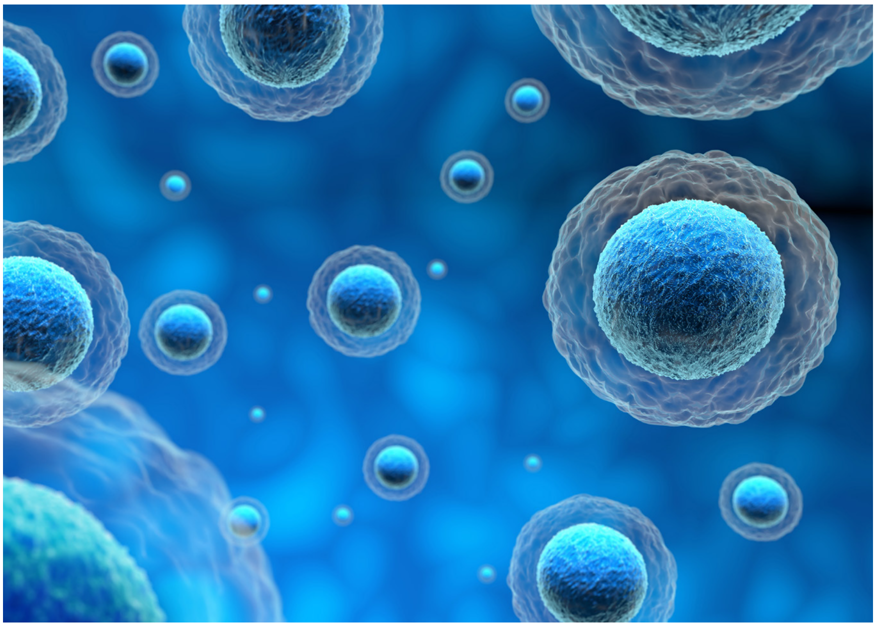
8- Vamathevan, J., Clark, D., Czodrowski, P. et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18, 463–477 (2019). https://doi.org/10.1038/s41573-019-0024-5

9- A. Avati, K. Jung, S. Harman, L. Downing, A. Ng, and N. H. Shah, "Improving palliative care with deep learning," 2017, pp. 311–316, doi: 10.1109/bibm.2017.8217669 [Online]. Available: https://ai.stanford.edu/~avati/bibm17.pdf

10- Flory JD, Yehuda R. Comorbidity between post-traumatic stress disorder and major depressive disorder: alternative explanations and treatment considerations. Dialogues Clin Neurosci. 2015;17(2):141-150. doi:10.31887/DCNS.2015.17.2/jflory



An example of our algorithm's prediction of patient outcome. Blue line is the algorithm's prediction. Yellow dotted line is the actual MADRS score. Vertical magenta line indicates the 60-day mark. To the left of the magenta line the system was trained on actual data. To the right of the magenta line the system is projecting patient outcomes based on physiological data collected from a wearable device. Our algorithm correlates very closely with ground-truth MADRS, and is capable of operating far beyond human horizons of prediction.



Ketogenic diets increase the amount of ketones in the blood, which reduces the oxygen demand for certain cells. Could this popular diet actually help other damaged cells and tissues repair faster? This basic research explores the science of the ketogenic diet and its effects on cellular oxygen demand. (Stock photo)

PERFORMANCE THROUGH DIET

Taking a close look at the ketogenic diet and how the effects of cellular performance could mean big gains to Naval Aviators

By: LT Claire Modica, PhD, MSC, USN, AEP #157

Original research was conducted at Naval Medical Research Center by LT Claire M Modica, Krystal Flores-Felix, LT John D Casachahua, Paul Asquith, Anna Tschiffely, Stephanie Ciarlone, and Stephen T Ahlers

There are circumstances among military stressors where cellular access to oxygenation may be altered. In aviation and high-altitude environments, partial pressure of oxygen is decreased while in undersea environments, partial pressure of oxygen is increased. In contusion or hemorrhage injuries, when blood vessels are ruptured, some tissue may become inundated with oxygenated blood while other tissue experiences reduced oxygenated blood flow. Even, simply, in times of elevated heart rate or cardiovascular activity, tissue demands for oxygenation can exceed supply. This leads military researchers to pose the question: Would performance under these circumstances differ if the metabolic need for oxygen is decreased?

One method under which to explore altered metabolism is to use the ketogenic diet. In neurological trauma research, recent literature indicates that cellular production of energy decreases after injury. This has led researchers to explore ketone body metabolism, under which an oxygen-heavy process called glycolysis is skipped. This might free up oxygen availability for cells under inflammatory and repair processes. This is most easily done by utilizing the popular ketogenic diet as a tool. In the ketogenic diet, circulating ketone bodies are increased by decreasing carbohydrates and increasing the proportion of fat in the diet. Fatty acids are broken down into ketone bodies which then get turned into ketone esters, and ketone esters are turned into a substrate for the Citric Acid Cycle, where a large proportion of cellular energy is produced.

Scientists at Naval Medical Research Center are exploring how the ketogenic diet, or ketone ester supplementation, affects injury and recovery from blast exposure. In a 2021 Food Chemistry: Molecular Sciences article, they characterized the diet models established in their lab (Modica et al., 2021). The publication detailed the daily body weight, blood glucose concentration, and blood ketone concentration in male and female rats over the first two weeks of changeover from a standard lab diet. They found that both the ketogenic diet and the ketone ester supplementation increased blood ketone concentration and prevented age-related weight gain that was otherwise expected on the standard diet. Overall, the effects of the diet were larger than the effects of the supplement. In addition, the diet had an effect on glucose in males. It was observed that the glucose concentration in males on the standard diet was relatively high compared to males on the ketogenic diet and high compared to females on all diets. The prevention of weight gain occurring simultaneously with decreased glucose levels upon diet changeover suggests a possible pre-diabetic condition on the standard diet, and the difference in effects between the sexes indicates that males might be particularly susceptible.

An East-Coast based U.S. Navy SEAL climbs a caving ladder during visit, board, search and seizure (VBSS) training on Joint Expeditionary Base Little Creek-Fort Story, July 16. Navy SEALs and others are at high risk for blast injuries due to the nature of their role in maritime safety and security. Research such as this study may lead to interventions that can help speed recovery from such injuries at the cellular level. (U.S. Navy Photograph by Mass Communication Specialist 2nd Class William S. Parker/Released)

Possibly the most interesting observation in this paper was that ketone levels surged within the first few days after changeover to the ketogenic diet prior to coming down nearly (but not entirely) to standard diet levels. In contrast, this was not observed in the ketone ester supplement; in the supplement, ketone levels increased mildly and remained constant. The decrease in ketone levels observed in the diet condition may reflect a state of fat-adaptation, or as it is sometimes called, keto-adaptation. This generally describes a state in which, even upon an increase in carbohydrate intake or a decrease in dietary fat, the body continues to metabolize ketone esters. It is not yet understood, but it is possible that fat adaptation is a condition in which enzymes responsible for breaking down fatty acids, ketone bodies, and ketone esters are upregulated, increasing the likelihood of fat breakdown. Glycogen stores might even be present and preserved in this state, while ketone esters are outcompeting glucose in providing substrate for the Citric Acid Cycle. Older literature on fat adaptation indicates that not only does the ketogenic diet bring people and animals into a fat adapted state, but mildly elevating the proportion of fat in the diet over months will do so as well, suggesting that there are many ways to achieve a fat adapted state, given enough time. While ten days may be enough time for rats to become fat adapted on the ketogenic diet, perhaps it takes more than two weeks for the adaptation on the ketone ester supplement. Furthermore, bouts of fasting, which also triggers the breakdown of fatty acids (but autocannibalistically of one's own fat rather than external dietary fat), might lead to fat adaptation if done long enough or frequently enough. Even long-term endurance athletics, during which fatty acids are broken down and metabolized, could lead to some degree of fat adaptation.

If ketone metabolism leads to improved repair and recovery, then a fascinating notion is that a combination of endurance athletics and fat adaptation could spiral into performance optimization. For example: fat adaptation might have an accelerating effect on neuromuscular recovery, which could facilitate more frequent pushes in athletic training. That could create an increased caloric need during training, resulting in further autocannibalism of the body's fatty acids. Longer duration in periods of fat breakdown may further upregulate genes required for such breakdown, leading to a greater degree of fat adaptation. In advanced states of fat adaptation, when it is possible that fat, fatty acid, ketone ester, and ketone enzymes are in abundant supply, it would be imaginable that fat would continue to be the primary source for energy, even while glucose is available. It may even be the case that glucose simultaneously gets broken down to fuel the Citric Acid Cycle for bouts of excitatory neurotransmission for some parts of the body while other parts continue to be fueled by ketones. These are testable hypotheses, and further research on these aspects of human performance optimization are sure to follow in the coming years.

CRACKING THE CODE OF BLAST INJURIES

Finding reliable methods of estimating prior blast exposure may yield critical insights for treatment of traumatic brain injury

By: LT Claire M. Modica, PhD, MSC, USN, AEP#157

Original research was conducted at Naval Medical Research Center by LT Claire M Modica, Michael J Egnoto, Jonathan K Statz, Walter Carr, and Stephen T Ahlers

In the absence of blast wave overpressure sensors, there has been no method for standardized characterization of prior blast exposure. LT Claire Modica, AEP #157, and her team published a method of blast exposure estimation in November 2020 in Journal of Neurotrauma (http://doi.org/10.1089/neu.2020.7405). The Blast Exposure Estimator consists of a two-step process by which to represent an individual's lifelong blast exposure with a single number. The first step requires an individual to respond to the Categorization of Light arms, Artillery, Recoilless rifles, and Explosives (CLARE) Query. This Query design was informed by service member feedback from pilot testing: volunteers were asked to group weapons according to which blasts were more similar versus more dissimilar from one another. The second step consists of plugging the Query responses into the Generalized Blast Exposure Value (GBEV) Formula. The Formula combines numerical responses from the Query with data-driven weighting factors derived from an anonymous survey study of nearly 1000 service members. At the end of the two steps, the resulting GBEV characterizes all blast exposure experienced by an individual in a single number. Since the process takes about 5 minutes, this number can

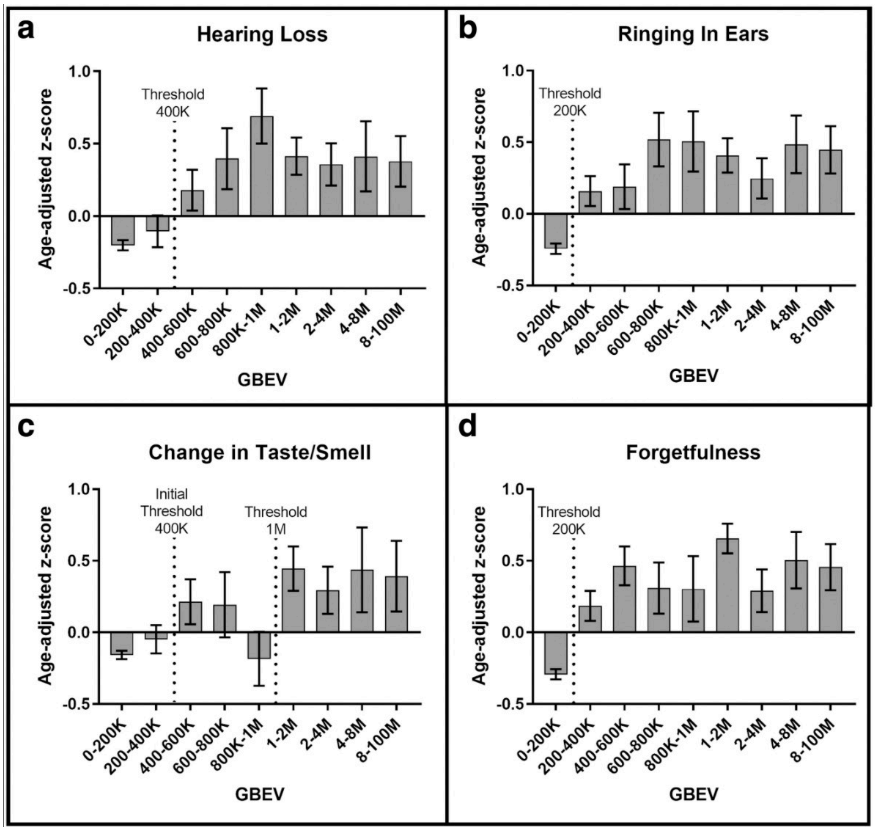


Figure 1. Age adjusted z-scores of blast-associated outcome reporting intensity for (a) hearing loss, (b) ringing in ears, (c) change in taste/smell, and (d) forgetfulness. Data points were first binned from 0-200,000, then subsequent bins were structured not to exceed 100 data points per bin. Thresholds drawn to represent the GBEV unit at which outcome intensity extends above mean. Error bars represent standard error. GBEV, generalized blast exposure value; K, thousands of GBEV units; M, millions of GBEV units. Source: Journal of Neurotrauma. Copyright 2021, Mary Ann Liebert, Inc., publishers

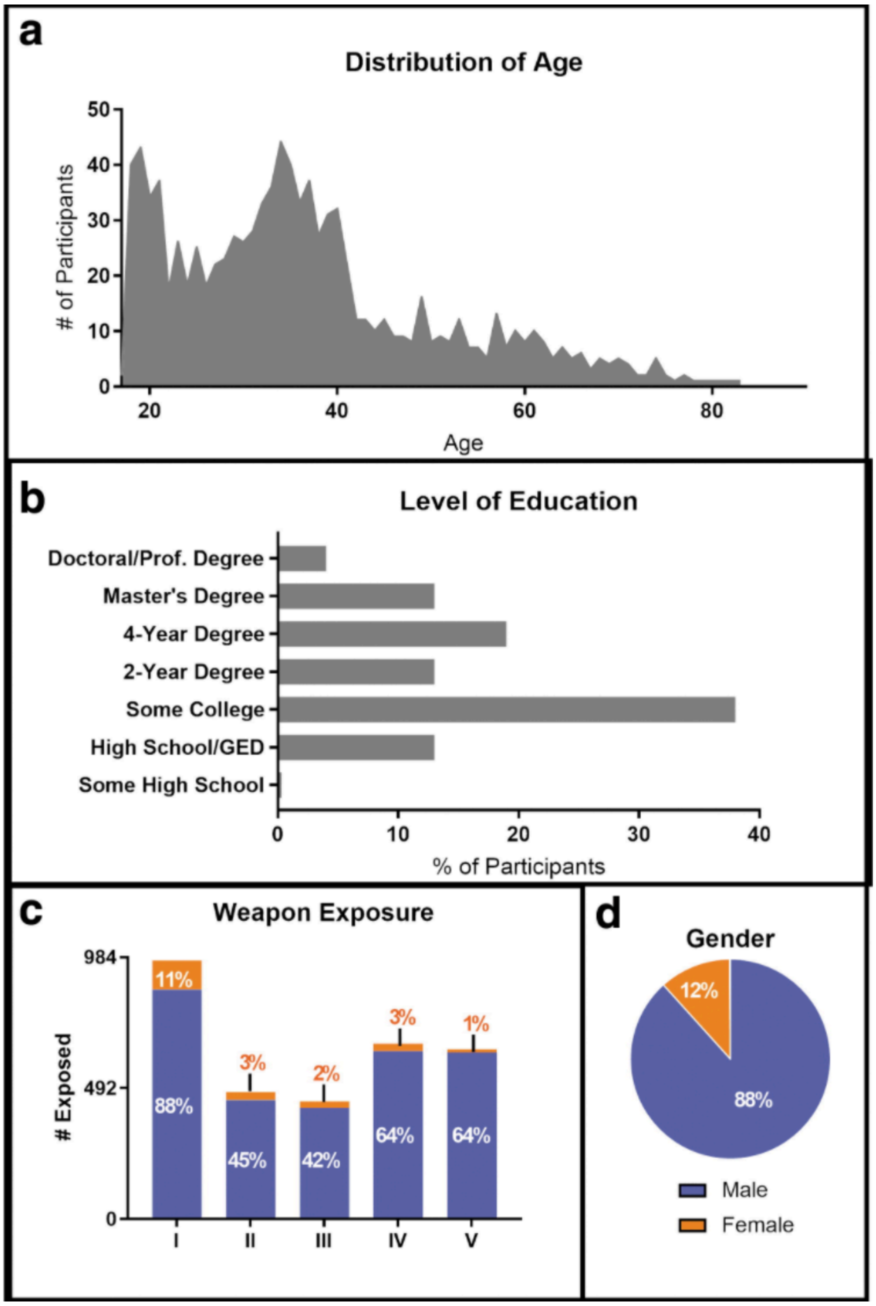


Figure 2. Participant composition: (a) age distribution; (b) level of education; (c) number of participants exposed to each weapon category, broken down by men and women in stacked bars with percentage overlaid; (d) gender. Categories = I: Small Arms; II: Large Arms; III: Artillery; IV: Small Explosives; V: Large Explosives. Source: Journal of Neurotrauma. Copyright 2021, Mary Ann Liebert, Inc., publishers

be added proactively by filling out a new CLARE Query, then plugging into the GBEV Formula, as frequently as necessary.

$$\begin{aligned} GBEV &= 0.976(1BEC) + 288(2BEC) + 41(3BEC) \\ &\quad + 77(4BEC)(4freq) + 75(5BEC)(5freq) \end{aligned}$$

For example, using the equation to the right, we can calculate a 30-year-old man with a high school/GED degree, 18 years of small arms, 1 year of large arms, and 1 year of small explosives exposure as having 1034 GBEV units; a 34-year-old man with a 4-year degree, 10 years of small arms, 2 years of large arms, 1 year of artillery, 8 years of small explosives, and 8 years of large explosives exposure as having 100,031 GBEV units; and a 32-year-old man with some college, 14 years of small arms, 14 years of artillery, 14 years of small explosives, and 1 year of large explosives exposure as having 1,181,195 GBEV units.

Not only can the GBEV be used to tally the accumulation of blast exposure over time for record-keeping, it could be used as a screening mechanism. At a threshold of approximately 200,000 GBEV units, some individuals begin to exhibit a greater degree of outcomes compared to similarly-aged service members. Being associative, and not causal, these outcomes are not necessarily due to blast overpressure, but could have to do with concomitant experiences or exposures. Nevertheless, the GBEV threshold can behave as a referral signal for further evaluation of service member health. The GBEV will also serve as a valuable covariate or independent variable in military research going forward.

Left to right: LCDR Micah Kinney, Aerospace Optometrist; CDR (ret) Mike Reddix and LT Sarah Sherwood, Aerospace Experimental Psychologists. The Green-X research team prepares to embark on an H60 Jayhawk helicopter with the US Coast Guard as part of testing and evaluation of laser eye protection.

LASER SAFETY FOR AVIATORS

Combating emerging threats to aviation safety through collaborative research to develop technological solutions

By: LT Sarah Sherwood, LCDR Brennan Cox, and CDR (ret) Mike Reddix

Laser illumination of military and civil aircraft, both in the continental U.S. and abroad, poses a threat to aircrew safety, performance, and mission effectiveness. Between 2007 and 2019, the Federal Aviation Administration observed a nine-fold increase in cockpit laser-illumination events (i.e., laser strikes), with 5,486 events recorded in CY19 alone. Between CY17 and CY18, the U.S. Coast Guard reported 75 lasing events, the U.S. Air Force reported 81 lasing events, and the U.S. Navy reported 127 lasing events. These events introduce a variety of safety and performance challenges, including aviator distraction, loss of situation awareness, and temporary scotomas (i.e., flash blindness), which, in turn, lead to reduced reaction times, obscured in- and out-of-cockpit information, and improper flight control inputs.

The safety threat posed by laser strikes is magnified during critical phases of flight, such as take-off, landing, and low-level operations. US Coast Guard air assets, which are slow-moving rotary-wing and fixed-wing platforms that operate near populated coastal areas, are particularly vulnerable to hand-held lasers. To address this threat, the Coast Guard requested a low-cost, nighttime capable low intensity threat laser eye protection (LIT-LEP) spectacle to mitigate the laser veiling-glare (i.e., laser dazzle) and temporary scotoma effects produced by hand-held lasers. To be acceptable, the LIT-LEP solution would need to be compatible with cockpit instrumentation, night vision devices, head-up displays (HUDs), and out-of-cockpit visual aids.

The Green-X research team, led by retired CDR (ret) Mike Reddix at Naval Medical Research Unit – Dayton (NAMRU-D), successfully executed a requirements-to-acquisition response to this high priority program requirement. The Green-X team was a joint-service and industry partnership that leveraged over $9M in prior Air Force Research Laboratory dye-polycarbonate LEP research and development and no-cost industry avionics subject matter expert support. The team proactively engaged Coast Guard senior aviation leadership to refine requirements, explore preliminary low-cost LIT-LEP spectacle solutions, and identify potential research sponsor interest. Preliminary material solutions were evaluated by Coast Guard aviators with the assistance of NAMRU-D and the Air Force Research Laboratory. Follow-on refinement of LIT-LEP requirements and potential solutions improved the operational focus of research proposals and led to the timely award of Defense Health Programs research funding.

The Green-X research team completed multiple Research, Development, Test, and Evaluation activities, including developing a low cost per unit LIT-LEP prototype development by the Air Force Research Laboratory's Materials and Manufacturing Directorate, Photonics Materials Branch; a spectral analysis of all six Coast Guard fixed and rotary-wing aircraft cockpits; laboratory psychophysical testing while wearing LIT-LEP prototypes; night vision device interoperability; static and dynamic flight simulator evaluations in support of flight safety; design and coordination of flight test evaluations; and preliminary Coast Guard Aviation Life Support Equipment acquisition support.

For the laboratory studies, a repeated-measures experimental design was used to assess LIT-LEP performance relative to a no-LEP control for the following tasks: near- and far-contrast acuity, night vision far-contrast acuity, emissive and non-emissive light source color-vision screening, and Coast Guard multifunctional display color symbol discrimination reaction time and accuracy. Near- and far-contrast acuity results demonstrated good LIT-LEP performance for typical in- and out-of-cockpit lighting conditions. Night vision device performance suffered marginally at only one contrast level (85%; 20/30 acuity line). Color vision test results showed good color balance in that S-, M-, and L-cone performance did not demonstrate a clinical diagnostic color defect for emissive or non-emissive light sources when wearing LIT-LEP. Color symbol discrimination reaction-time-task results based on inverse efficiency scores revealed that some non-primary flight display colors exhibited a combination of slower speed and decreased accuracy.

To transition Green-X from the laboratory to the field, independent engineering evaluations were conducted in partnership with Naval Aviation Systems Command to support flight safety. Close coordination with USCG CG-711, Coast Guard Aviation Training Center, and Coast Guard aviation life support equipment led to expedited ground evaluations by 13 aircrew and simulator flights by 24 aviators. The process culminated in over 21 hours of flight testing by 14 aviators and 7 critical position aircrew members.

NAMRU-D uniformed aircrew scientists, Aerospace Optometrist LCDR Micah Kinney and Aerospace Experimental Psychologist LT Sarah Sherwood, were by-name requested to guide USCG flight test development and to support safety of flight as onboard observers for flight-testing across all six Coast Guard fixed- and rotary-wing aviation platforms (HC-130H, HC-130J, HC-27, HC144A, MH-60T, and MH65D). Fixed-wing flight test events required the aviators to perform a normal start-up to ensure they could see engine Torque, Turbine Temperature, NP (propeller rotation speed), and NG (gas generator speed) indicator colors as they transitioned from red to yellow to green. Evaluated flight tasks included normal takeoff, instrument landing system and visual approaches, low-level flight over water, cruising at altitude, and patterns. Operationally relevant visual tasks included identification of maritime vessels, objects in the water, spotting aircraft traffic, and correctly calling the location of lighted buoys and navigational aids. In addition, aviators confirmed they could see runway and taxiway lighting, thresholds, PAPI/VASI glideslope indicators, and other airfield colored lighting.

Similar to the fixed-wing flight events, rotary-wing flight events required the aviators to perform a normal start-up to ensure they could see all engine information. Visual compatibility checks were performed for all displays and lighting (both in- and out-of-cockpit) to include combined wear of the LEP spectacles and AN/AVS-9 Class B NVGs with light-interference filters installed. The team employed surveys and open-ended questions to evaluate LEP spectacle fit and comfort when worn alone and during combined wear with helmets and NVGs. The team and USCG aviators and aircrew evaluated LEP spectacle performance during a representative set of flight profiles and mission tasks, to include: hovering over land and water, ILS approach, visual approach, takeoff, departure, patterns, manual approach to a controlled hover over water, approach to a maritime vessel, hoisting of a mannequin and rescue swimmer, and delivery of rescue supplies to a vessel. Color perception was evaluated around navigational aids and airfield environments.

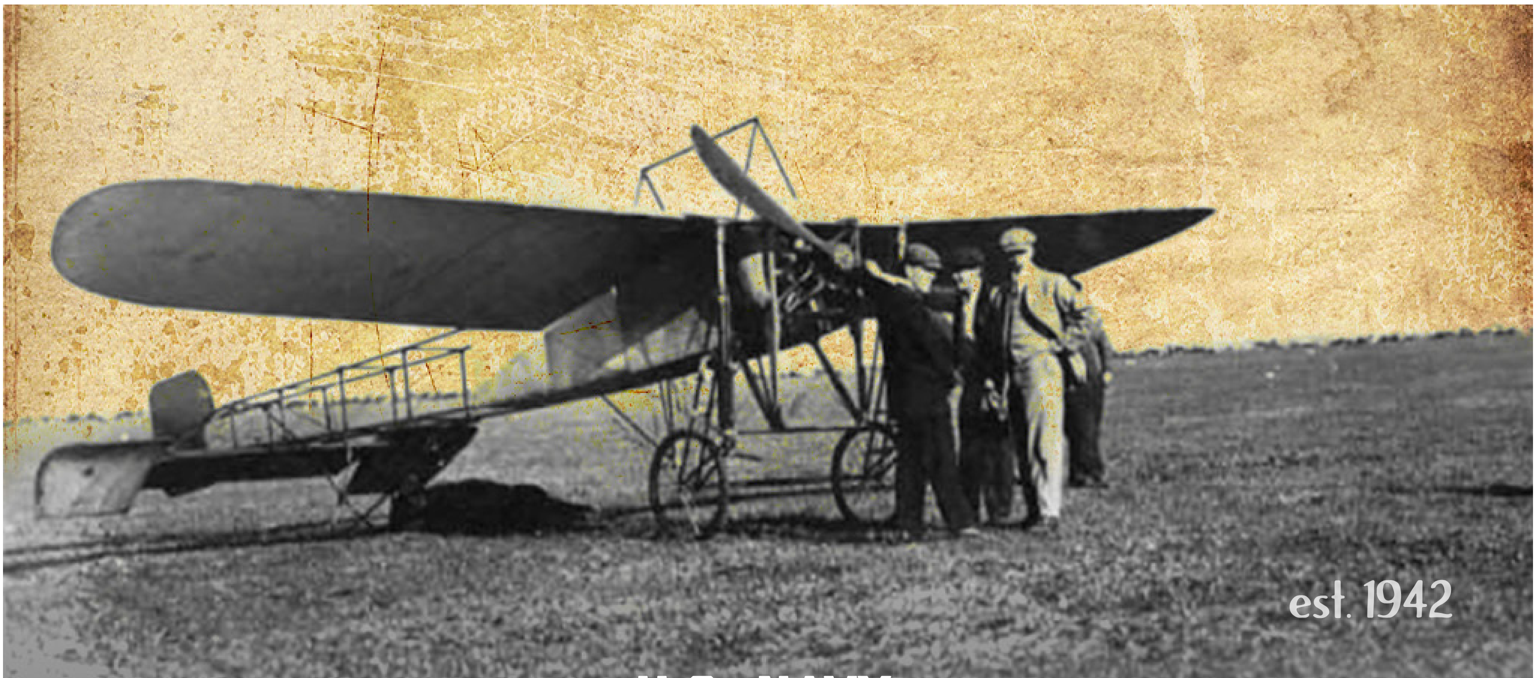
These low-cost laser safety spectacles were developed in partnership between AEPs and scientists in the US Air Force.

During all test flights, the team collected survey data on LEP spectacle performance from aviators and critical position aircrew as it related to the: In-cockpit environment; Out-of-cockpit environment; Over land environment; Over water environment; and Viewing of critical instruments. Sections for overall assessment and open-ended responses were included for any feedback, opinions, or issues that may not have otherwise been captured by the survey. Members of the research team also captured in-flight verbal statements and comments and asked clarifying questions as needed.

The results of the Green-X studies and user feedback were highly favorable in both compatibility and visual performance. This partnership with the Air Force Research Laboratory and NAMRU-D resulted in a LIT-LEP technical specification publication, four technical reports, a peer-reviewed psychophysical evaluation publication, and expedited the modeling of future prototype LIT-LEP designs for cockpit compatibility. Flight testing directly informed the USCG ASLE recommendation to acquire LIT-LEP. An acquisition partner was identified and the fully transitioned product, an LEP spectacle that is compatible with 6 USCG aircraft platforms, is available for purchase via GSA. This remarkable collaboration between NAMRU-D, AFRL/RXAP, Collins Aerospace, and USCG senior aviation leadership resulted in a series of laboratory and live-flight studies to transition research from Advanced Technology Development (6.3) to a scientifically supported fielded, operational product in just five years, directly addressing a fleet need for combating an emerging airspace threat.

LT Sarah Sherwood (left) and LCDR Micah Kinney prepare for a flight on board a US Coast Guard H-60 as part of their research into mitigating laser strikes on aviators.

CDR (et) Mike Reddix, a former AEP and current research scientist at NAMRU-Dayton sits behind the polycarbonate glass of a mock cockpit to demonstrate and test the laser-resistent eyeware they developed. This project is an example of the 6.3 budget activity on the RDT&E spectrum.



AUTONOMOUS CASUALTY EVACUATION

Demonstrating capabilities and evaluating results is the cornerstone of 6.5 activities on the RDT&E spectrum.

By: LT E.S. Vorm, PhD
Deputy Director, Laboratory for Autonomous Systems Research
US Naval Research Laboratory, Washington, DC

(Left) Marines assault an enemy position as part of their training scenario. (Below) A team of Marines provides cover for wounded casualties as they are evacuated via our autonomous vehicle out of the firefight to be treated at the Battalion Aid Station.

(Above) The Hunter WOLF, developed by HDT Expeditionary Systems, is prepared for arduous testing in real-world conditions at Range 220 onboard US Marine Corps Base Twentynine Palms. This project was created and led by LT ES Vorm, PhD from the US Naval Research Laboratory's Navy Center for Applied Research in Artificial Intelligence as part of a 6.5 funding activity to demonstrate and evaluate the concept of autonomous casualty evacuation.

(Above) A Marine walks beside the Hunter Wolf as his platoon patrols the open area en route to their objective. The Hunter WOLF is configured here to carry four casualties; two in the center and two on the wings.

In the barren Mojave desert, not far from Joshua Tree National Park, a mock battle rages. Commanders shout orders. The clatter of machine gun fire rings out and echoes off nearby buildings. A squad of Marines presses their bodies against the dirt on the lower edge of a berm and try to return fire against an unknown and unseen adversary. Suddenly a Marine flinches and turns over, yelling for help.

“Corpsman up!” screams another Marine.

Shots ring out from all directions. The air is filled with the sharp tang of gunpowder. Dust wafts through the scene as more Marines dash across the sandy ground, seeking cover. The din of radio chatter mixes with the shouts and cries of the injured Marine who lies writhing on the ground, grasping at his leg in pain.

Twenty-five-year-old Hospital Corpsman Second Class (HM2) Curtis Ikkala, a Fleet Marine Force corpsman, arrives on the scene and begins to treat the wounded.

“It is more than a half-mile to the CCP,” he says, referencing the casualty collection point—an intermediate collection spot where wounded Marines can be

The research team consisted of uniformed and civilian scientists and engineers from both industry and the DoD. Right to left: LT E.S. Vorm, PhD, primary investigator; LCDR Brennan Cox, PhD, LT Sarah Sherwood, PhD. Not pictured: LtCol Bryan Patterson, USMC; Kent Massey, HDT Expeditionary Systems; Charlie Deaver, HDT Expeditionary Systems; Michael Hodgson, San Diego State University

treated away from the most intense fighting. “This is gonna suck.”

The Marines maneuver the injured patient onto a litter, then lift him to waist-height and prepare to make the journey overland. Four Marines, one on each corner of the litter, carry the patient and his gear—approximately 240 pounds—while two others and the corpsman provide cover.

These Marines are with the Second Battalion, Fifth Marine Regiment out of Camp Pendleton, California. They have come out to the vast desert area of Twentynine Palms for a month of intense, realistic training. Today’s exercise is specifically focused on casualty evacuation—the ability to coordinate the movement of injured people to higher echelons of care while maintaining tactical superiority.

> **“IF WE HAD TO DO THIS IN REAL LIFE, WE’D IN BE A BAD STATE REALLY QUICKLY.”**

The unfolding scene looks very familiar to anyone who has experienced ground combat operations over the past decade: difficult, cluttered terrain; multiple moving groups of people; chaotic communications; intersecting fields of fire; limited sight lines; and multiple layers of concealment all combine to make the movement of wounded patients a slow, painstaking process. And there is something else about this scene that is also oddly familiar: the sight of multiple Marines carrying one wounded patient. Although today’s Marines benefit from superior weapons and technology, the process of moving patients from the point of injury to higher echelons of care has barely changed from methods used hundreds of years ago.

The group maneuvers down a narrow alleyway and descends a steep, sandy berm. The Marines holding the litter struggle to manage under the intense weight. Their hands throb and ache. Their movements are jarring and random as the litter carriers jostle and bump against each other in constant motion over the loose, uneven terrain. Sweat pours from their faces. Their heavy, labored breathing is punctuated by the moans of the patient, who may be just role playing, or may genuinely be complaining because of the roughness of the ride; it is impossible to tell.

Suddenly, one of the Marines loses his grip on the litter. His corner of the litter drops, which quickly cascades into the entire litter falling to the ground. The patient bounces violently and rolls partially off the litter, disrupting the sensitive medical interventions--a tourniquet, two pressure dressings, and an IV--that have thus far kept him notionally alive.

In tactical situations such as this, where large vehicles such as HMMWVs are unavailable, carrying patients from one place to another presents a crude, but mostly effective solution. Military conflicts from as far back as the ancient Romans have featured some form of carrying device (i.e., a litter). The American Civil War in the 1860s introduced litter carriers--teams of people that were specifically designated to accomplish this task. World Wars I and II also saw these roles expanded to include more dedicated medical personnel with equipment such as jeeps and field ambulances, and tactics that enabled the strategic and organized movement of patients. The Vietnam war introduced the concept of air ambulances for patient movement, and forward aid stations that acted as intermediate patient collection points to stabilize patients with limited surgical interventions. These casualty evacuation concepts continued to evolve through conflicts in Kosovo and Desert Storm in the 1990s, through more recent conflicts in Iraq and Afghanistan in the 2000s and 2010s. Throughout these conflicts, the case fatality rate (CFR)--which measures the total lethality of the battlefield—has fallen precipitously from a high of 55 in World War II, down to 12 for the Iraq and Afghanistan conflicts. Many factors, such as mobile IV fluids and clot-enhancing pressure dressings; improved medical training for every warfighter; and better body armor, have contributed to this improvement in survivability, but one constant has remained—the humble litter.

“If we had to do this in real life, we’d be in a bad state really quickly” says HM2 Ikkala as he readies the patient and prepares to resume the long trek out of the hot zone back to the casualty collection point.

His words echo the concerns of many who study the current state of how the US military goes about the business of evacuating and treating patients on the battlefield. The challenge he is specifically referencing is the fact that one wounded Marine has taken an additional six Marines out of the fight—four to carry the patient, and two more to protect the group as they maneuver. This ratio is a troubling one in light of speculations of what tomorrow’s war against a near-peer adversary would mean for the United States.

Casualty projections in the event of a kinetic fight with a near-peer adversary such as China, based on expert analysis and war gaming, indicate grim statistics: U.S. forces are projected to experience orders of magnitude more casualties than anything the past generation has had to face. The current manpower-to-patient ratio means that a platoon-sized element, roughly 50 people, could only sustain 4-5 casualties before being overwhelmed by the logistical burden of treating and maneuvering patients out of the fight. And the challenges for expeditionary medicine don’t end there.

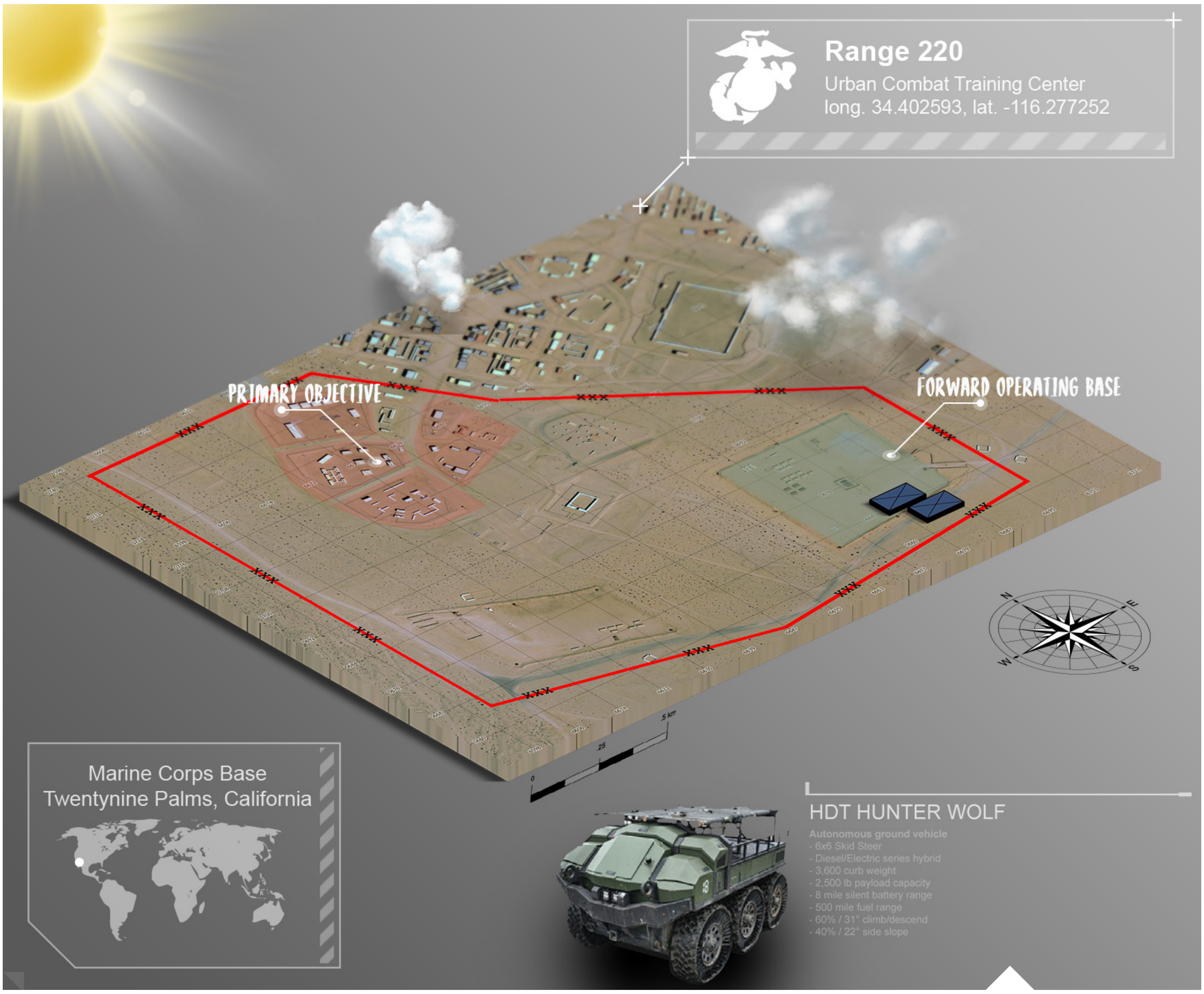
Tomorrow’s battlefields are expected to feature significant use of the electromagnetic spectrum in the form of jamming satellite communications and spoofing radar. This will force a radical change from how the US has coordinated its forces in the past by using large, centralized forces such as a carrier strike group or forward expeditionary force. Instead, units will need to operate more independently from one another, and will be potentially distributed across wide areas. With vast areas of the war zone blanketed in communications-degrading static and GPS-disrupting attacks, units may not have the freedom of movement they have enjoyed in recent conflicts—such as the use of helicopters or convoys of trucks to expeditiously evacuate patients who need advanced medical treatment.

The result of these factors paints a challenging picture for a military force that has grown accustomed to being able to move about freely in contested areas to deliver medical care to those who are injured. Through this speculative lens, the ratio of six Marines for every wounded casualty and the idea of carrying patients on collapsible litters becomes a critical problem in need of an innovative solution.

One such potential solution envisions the use of small, unmanned ground vehicles that are capable of traveling alongside dismounted infantry troops and can fill a variety of roles: logistics, supply, reconnaissance, and even casualty evacuation. Vehicles roughly the size of golf carts, capable of a full range of autonomy, could fill a niche role for isolated and distributed forces in difficult terrain where full-size vehicles like the HMMWV and others may not be able to go. The Hunter WOLF, designed and built by HDT Expeditionary Systems, Inc., was created for this exact scenario.

The WOLF is a six-wheeled vehicle that measures around 7.5 feet long, 4.5 feet wide, and just under 4 feet tall. It weighs around 3,600 pounds, which is around 30% lighter than the HMMWV, but can carry the same payload of 2,500 pounds, allowing it to carry military equipment like weapons and ammunition, weeks’ worth of food rations, or 12 troops fully laden with gear. It has an internal diesel generator capable of outputting 15kW of power—enough to power an entire command operations center and all its associated components. But the WOLF is far more than merely a mobile power generator. Its low center of gravity and extreme torque-to-weight ratio means

LT E.S. Vorm interviews participants in the study while on the march after the conclusion of another simulated combat mission. Over the course of three days the research team conducted 18 individual trials and gathered data from 2,042 active duty participants. Their input is critical to the evaluation and refinement of the ACE concept.



For three days in October 2021 Marines from 2nd Battalion 5th Marine Regiment participated in extended combat training on Range 220 at Marine Corps Base Twentynine Palms. Scenarios dedicated to casualty evacuation enabled the research team to compare Marines' performance in evacuating casualties using traditional litters, and an autonomous ground vehicle built by HDT Expeditionary Systems called the Hunter WOLF.

that it can tow more than three times its own weight, while its small compact size and zero-turn radius means that it can maneuver through the narrowest alleyways with ease. It is fitted with innovative features like Michelin airless radial "Tweels" that conform to terrain and never go flat, and a hybrid diesel-electric motor that can go for hundreds of miles silently on battery power. The WOLF was designed to support a platoon-sized group of dismounted infantry for five days with no resupply; to move alongside troops and be adaptive to a variety of roles; and is an ideal solution for scenarios where regular and extensive supplies like fuel, ammunition, and food may not be available for days or even weeks.

I have come to this exercise in the desert training grounds of Twentynine Palms to evaluate how vehicles such as the WOLF can help accomplish a new, emerging concept in military medicine: autonomous casualty evacuation (ACE). In scenarios where operating rooms and extensive medical infrastructure are not available, units will need to be able to provide both critical, life-saving interventions and also sustain patients for prolonged periods of time under difficult conditions. Automation of monitoring and treatment is the cornerstone of the ACE concept, in which machines driven by artificial intelligence could lessen the burden on medical providers by autonomously monitoring patient vital signs and providing limited clinical interventions such as administering fluid resuscitation or medications. Other necessities such as IV warmers, powerful suction, and mechanical ventilation would enable medical providers a broad range of treatment options in a field environment. And of course, the most radical of the ACE concept involves using autonomous or semi-autonomous vehicles like the WOLF to transport patients, with or without a human "at the wheel."

Vehicles like the WOLF are obviously well-suited to provide the logistical means and power to integrate all of the monitoring and treatment technology necessary to accomplish the ACE concept in a small-enough footprint to remain viable in austere environments. More importantly, however, is the exponential value they could add by flipping the 6:1 ratio of current manpower requirements of casualty evacuation on its head. One WOLF could easily transport four patients on litters, and could potentially do so with a minimal amount of human supervision. Utilizing the same kinds of technologies that are enabling self-driving cars to enter public roadways today, vehicles like the WOLF could lessen the burden of transporting patients on the battlefield, resulting in more fighters staying in the fight.

To empirically demonstrate and evaluate the WOLF in a casualty evacuation role, we designed a simple factorial design. Each platoon in the battalion, one at a time, would execute the same objective of assaulting an area in order to capture and control a collection of buildings. Some groups would be given the WOLF as an asset to evacuate patients, while some groups would only be allowed to use traditional hand-carried litters. The traditional litter groups would serve as a control group against which the performance of the WOLF groups would be compared. This resulted in a two by five-way factorial design. Two experimental conditions: traditional litter carry for evacuation, and using the WOLF for evacuation were compared across five groups of people: drivers are those are have been designated to control the vehicle; security are those have been designated to provide physical security during the evacuation; corpsman are those who are designated to provide medical aid to the patient; patients are those who are designated with mock injuries; and for the traditional condition we have litter carriers, those individuals tasked with carrying the patient on a litter. Performance would be measured in two general ways: the subjective workload experienced by participants, and the quality and efficiency of teamwork as measured by structured observations.

To estimate workload, we used the NASA-developed Task Load Index (TLX), originally created to evaluate the workload of operators interacting with new spacecraft and robotics. Workload can be physical, as in the physical effort necessary to carry a patient on a litter, but also can be mental. For example, how much mental effort does it take to determine the best route the WOLF should take when moving patients? How difficult are its controls? How much time does it take to get the WOLF to do what you want it to do? How quickly can a person learn to control a vehicle like the WOLF, and at what point are they considered proficient? These are all branches off the tree of mental workload, and these questions are important to answer in order to ensure the WOLF fully meets the needs of the units it seeks to support.

Each platoon was allowed to develop their own plan of action (known in infantry terms as their scheme and maneuver). Each group chose to incorporate the WOLF in their own way, which included who would be designated to control it, where the vehicle would be stationed, and how it would move with the unit during their assault. The time it took for each platoon to execute their mission ranged from 1.5 to 3 hours, depending on the speed and efficiency of their coordination. My team would pay close attention to how each platoon conducted themselves, and would document decisions and actions in order to correlate those to each platoon's overall performance. Each day ran three platoons through the assault over three days' worth of testing, for a total of 9 evolutions; two using traditional litter evacuation, and seven using the WOLF.

A total of 2,042 Marines took part in our exercise. 116 Marines directly participated in the evacuation of casualties and chose to participate in our study. A breakdown of participants by their role is available in the table below.

To best understand how the WOLF improved or hindered each group's ability to evacuate their mock casualties, we examined measures of workload as measured by NASA TLX. We used independent samples t-test wherever appropriate to determine if workload differed significantly between the two groups. To augment these measures, we used structured observations with time-stamped photographs and notes. We also solicited feedback using a structured interview format.

Corpsman from 2nd platoon Golf Company of 2nd Battalion 5th Marines respond to a mock casualty while their platoon sergeant communicates with headquarters. Over this three-day evolution, Marines from 2/5 received dedicated training in casualty treatment and evacuation in urban combat scenarios in preparation for their upcoming deployment overseas.

MENTAL WORKLOAD

Both the WOLF group (M = 18.09, SD = 19.59) and the traditional litter carrying group (M = 13.28, SD = 11.43) reported relatively low mental workload for the task of evacuating patients. The differences between the two groups was not significant, t(114) = 1.64, p = .057, d = 15.8. It is worth noting, however, that the traditional litter carrying group reported slightly lower mental workload than the WOLF group, and there were notable outliers in the WOLF group, all of which were Drivers. This could be interpreted that operating the WOLF required higher mental workload for those operators, which makes logical sense. Each designated operator of the WOLF received approximately one hour of practice and instruction before taking part in the assault exercise. This was the minimum necessary time to ensure the safe and effective operation of the vehicle, but was evidently not enough time to eliminate the extra burdens that remotely controlling a full-size vehicle adds.

Changing the control interface, for example, may help alleviate some of the mental demands that operators of the WOLF reportedly experienced. For example, during the evolutions, we heard feedback from multiple operators who said that controlling the WOLF with the thumbstick was difficult to accomplish while wearing gloves. The relatively low force-feedback of the thumbstick made it very easy for the operator to over-torque the WOLF, which would cause it to lurch forward and could potentially create an unsafe condition. The WOLF's control interface is very intuitive to use, and provides extremely fine degrees of control, but with the addition of gloves and operators who are moving alongside the vehicle, many of those degrees are lost. All of this means operators have to think more carefully about what they are doing, which results in more mental workload.

While training would likely alleviate much of this mental demand, it bears consideration that operating vehicles like the WOLF inevitably add a degree of complexity to the equation of evacuating patients in combat situations, which can have an effect on the overall performance of the team and its mission. If vehicles like the WOLF will be operated or controlled by a single, designated operator, then this finding may not ultimately be that important as that person would likely receive adequate training and experience operating the vehicle.

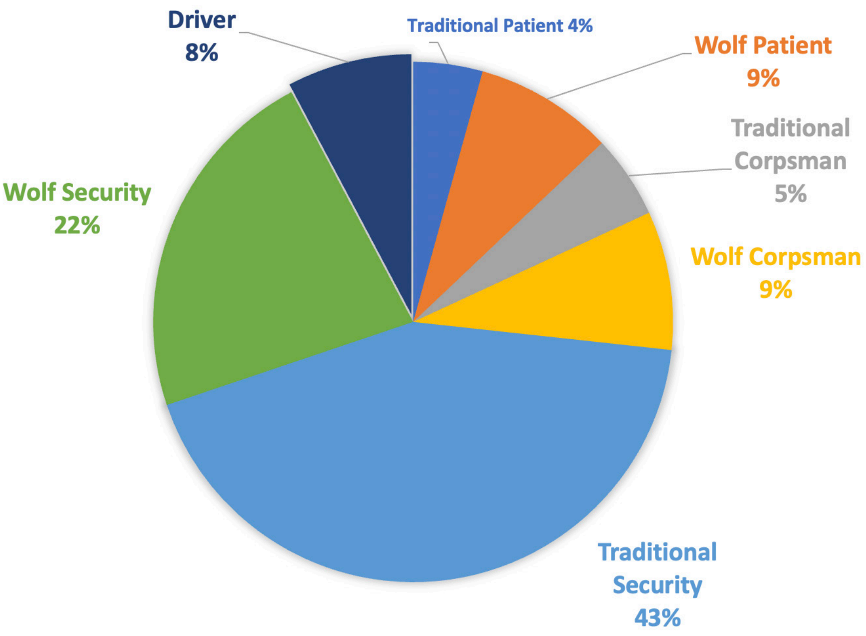


Figure 1: Demographics of participants in our study

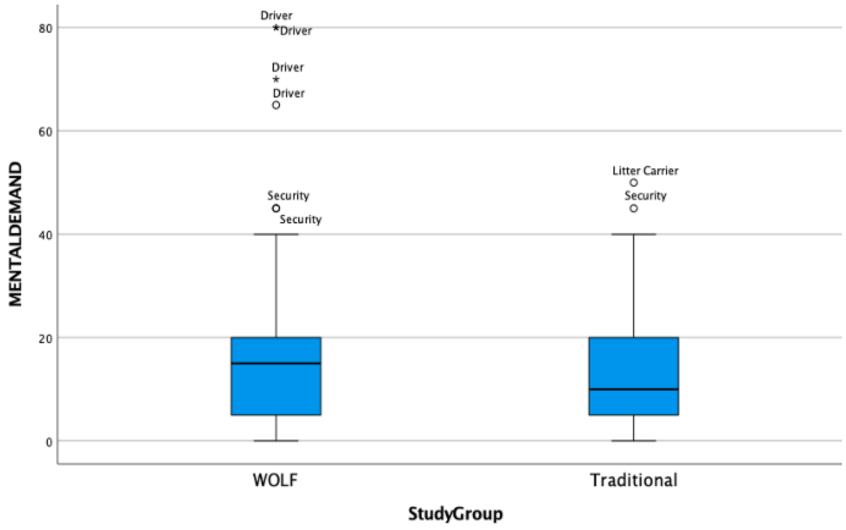


Figure 2: Box and whisker plot reporting the results of the mental demand experienced between two groups while evacuating casualties under battlefield conditions.

If the WOLF will be operated or controlled by multiple people, however, where some or all members of a unit have some cross-training but not necessarily extensive familiarity or experience with the system, then this finding becomes more critical. Findings like this help the design team to identify task and functional areas that can be improved.

PHYSICAL DEMAND

Physical demand was much higher for the traditional group than for the WOLF. Results showed that mean score for the WOLF group (M =13.2, SD = 19.8) was significantly lower than mean physical demand for those evacuating patients using the traditional litter method (M = 62.05, SD = 28.5), t(114) = 10.47, p < .001, d = 24.77. This result is not surprising and was predicted. The WOLF makes light work of carrying patients off the battlefield, while carrying them by hand is physically arduous. As before, we can see that while the overall average physical workload was lower for the WOLF groups over the traditional litter carrying groups, there are notable outliers. The single WOLF driver that reported 100 physical demand did so because he was the only person who had to respond to all of the casualties taken during his platoon's assault (there were four casualties for his exercise), whereas everyone else only had to evacuate a single patient before returning to the fight. The security participants that reported high physical demand explained their ratings as being related to lifting and positioning the patient onto the WOLF. We configured the WOLF to accommodate four patients on litters; two on top, and two on the sides of the vehicle, but the majority of the time patients were loaded onto the top of the vehicle. This means that a patient and all of their gear would have to be lifted approximately four feet high to be placed on top of the WOLF's bed. Even when spread across four people, this task takes a good deal of physical effort to accomplish. The personnel who were tasked with providing security tended to be the personnel who were responsible for loading and unloading the patient, hence their ratings of high physical workload.

Again, this data is valuable from a design improvement standpoint. A number of potential solutions could be developed, for example, a modular ramp system

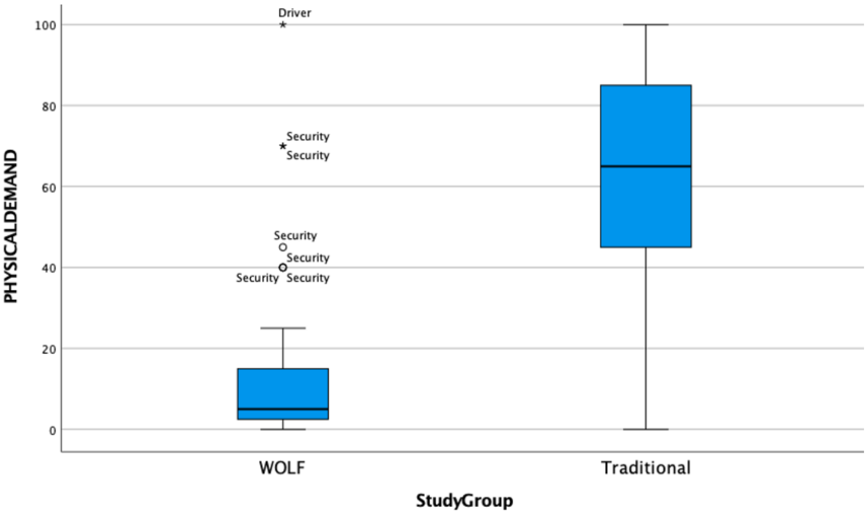


Figure 3: Physical demand was understandably much lower for the group using the Hunter WOLF, but there were some notable exceptions. These findings help to refine the WOLF's design and improve its usability for future operators.

that would enable a smoother and easier loading of patients onto the WOLF. By evaluating the WOLF under these realistic conditions with real operators, we can better evaluate and iterate on its design, which will result in a better overall product to meet the needs of our customers.

EFFORT

The perceived effort, as measured by NASA TLX was significantly different between the two groups, t(114) = 10.192, p < .001, d = 23.186. This means that the participants who were able to use the WOLF to evacuate their casualties thought the total effort involved, from physically moving the patient to transporting them to the casualty collection point, was less than those who had to perform the evacuation manually using a foldable litter. Effort can sometimes be thought of as a combination of physical and mental workload, in which case we see similar trends between our two groups reflecting perhaps the total effort it took to successfully evacuate patients using the WOLF versus using a traditional litter.

FRUSTRATION

The NASA TLX defines frustration as how insecure, discouraged, irritated, stressed, or annoyed a participant was when trying to accomplish the task. The WOLF group expressed significantly less frustration (M = 14.91, SD = 14.83) than the traditional litter carrying group (M = 55.41, SD = 22.12), t(114) = 11.45, p <.001, d = 19.01.

One way we can evaluate the differences in overall effort and frustration is to consider how many of the decisions made by different platoon leaders contributed to an easier or more difficult scenario involving the WOLF. For example, our team observed that many platoon commanders designated the lowest ranked person or a person who was not considered integral to any fire team to control the WOLF. Other platoon commanders did the opposite: they designated a staff sergeant or gunnery sergeant (E6 or E7) to control the WOLF. With higher rank also comes more experience leading troops, more authority to make decisions, and better decision making abilities. Thus, teams who used higher ranking people as operators had more coherent plans, communicated more efficiently, and executed their plans more successfully than teams with a very low ranking, less experienced individual at the controls.

Our team also observed a difference in how the platoons approached the WOLF as a strategic asset. Some platoons immediately saw the potential benefits of the WOLF, and worked to incorporate it into their scheme and maneuvers. These platoons used the WOLF in a variety of roles, expanding beyond only using it for casualty evacuation. For example, the WOLF was used to provide physical cover for moving troops on multiple occasions. It was used as a decoy and a distraction to fool enemy troops. It was also used to ferry supplies and people from location to location during the firefight, in terrain and under conditions that traditional vehicles would not have been able to afford. Conversely, some platoons appeared reluctant to use the WOLF and treated it as if it were a burden to them; some went so far as to leave it in the rear to wait until casualties were designated. These platoons appeared to think of the WOLF as a distraction to their overall mission, rather than a tool they could creatively use. Platoons who conceptualized the WOLF as a multipurpose asset and who saw it as a tool that could be creatively employed were more successful and required less

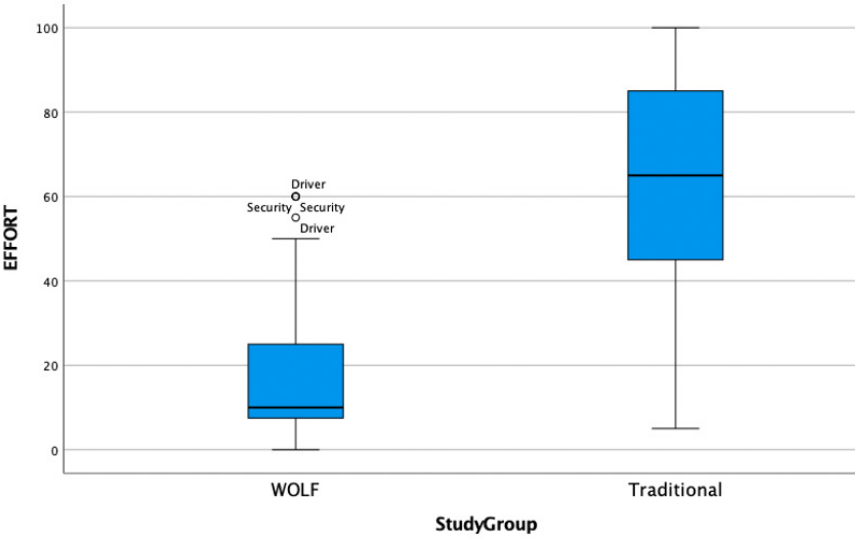


Figure 4: The overall effort it took to evacuate patients was lower for the Hunter WOLF group than using a traditional litter, but some roles experienced more effort than others.

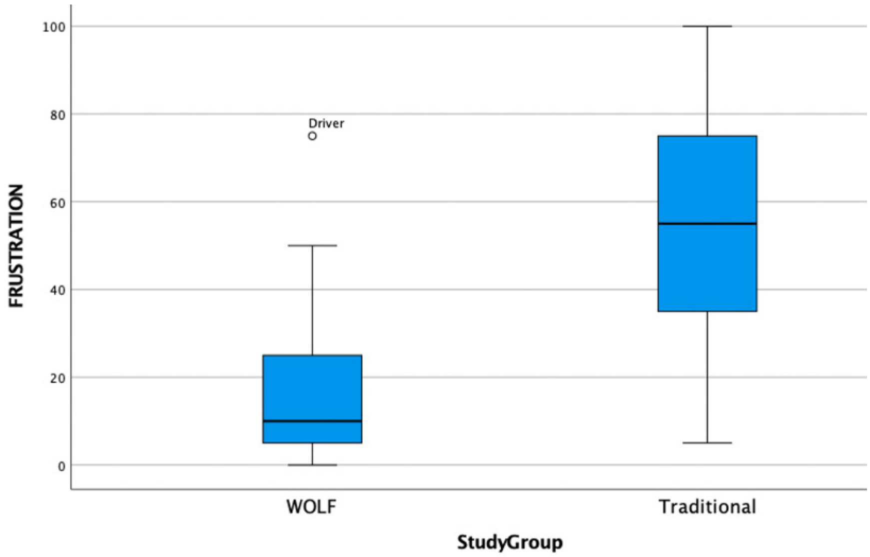


Figure 5: Frustration is a form of workload. Here we see the WOLF group had much less frustration than the traditional litter carrying group.

overall effort than other platoons.

This is an important finding if we consider a slightly broader scope. Technology is only useful when it is used, and used appropriately. How users think about and approach technology—seeing it through an adversarial lens or conceptualizing it as a teammate—makes a tremendous difference on the benefits that technology brings to bear. The overall effort these groups experienced was influenced, at least in part, by how they thought about and approached using the WOLF in their mission planning and execution. Future generations of Marines will no doubt be more familiar and comfortable with autonomous vehicles as they become more mainstream, but there will still need to be dedicated efforts to appropriately socialize these technologies in ways that engender trust and encourage them to be used. Failing to do so may result in technologies that ultimately hinder or slow performance, rather than speeding it up and making it easier.

TEMPORAL DEMAND

In dynamic situations with complex tasks, high physical and mental workload, and high levels of overall effort and frustration all tend to result in one thing: excessive time. The time it takes to accomplish a series of tasks is evaluated using NASA TLX, and is an important metric in our evaluation because when it comes to patients' lives, time is precious.

Due to the highly dynamic nature of each assault, we were unable to precisely measure how much time it took each group to evacuate patients from the point of injury to the casualty collection point. The perception of time as experienced by those involved in the exercise, however, was measured by NASA TLX. The participants' experiences differed significantly between the WOLF group (M = 18.2 SD = 18.9) and the traditional litter carrying group (M = 62.2, SD = 27), t(114) = 10.17, p < .001, d = 23.49. Along with being less physically challenging, participants using the WOLF for casualty evacuation appeared to experience quicker results. From our observations, it appeared that Marines using the WOLF for evacuation were indeed better able to move patients once they were loaded, but the coordination involved in moving the WOLF to the patient's location was sometimes slow and difficult, which sometimes slowed things down. Both the drivers and patients presented as outliers in these findings, which again makes sense in context. Both the drivers and patients needed to wait while patients were maneuvered over to the vehicle, and then again wait while they were loaded and unloaded. From the drivers' and patients' perspectives, these steps must have felt like they were taking a long time.

It is also worth noting that although physically carrying a patient on a litter may be tiring, it is a simple exercise—a physical action that all humans are familiar with—whereas maneuvering and controlling a vehicle from a third-person perspective and ensuring that patients are loaded appropriately so that the vehicle does not inadvertently dump them off if it turns too quickly are variables that few people are used to worrying about. It is important to examine these outliers in order to best understand the "pain points" of using vehicles like the WOLF. The best predictor of success in complex sociotechnical systems is the goodness of fit between the technology and its intended audience. Even the most sophisticated running shoe is of no benefit to a runner if it doesn't fit. Similarly, it is imperative that we design vehi-

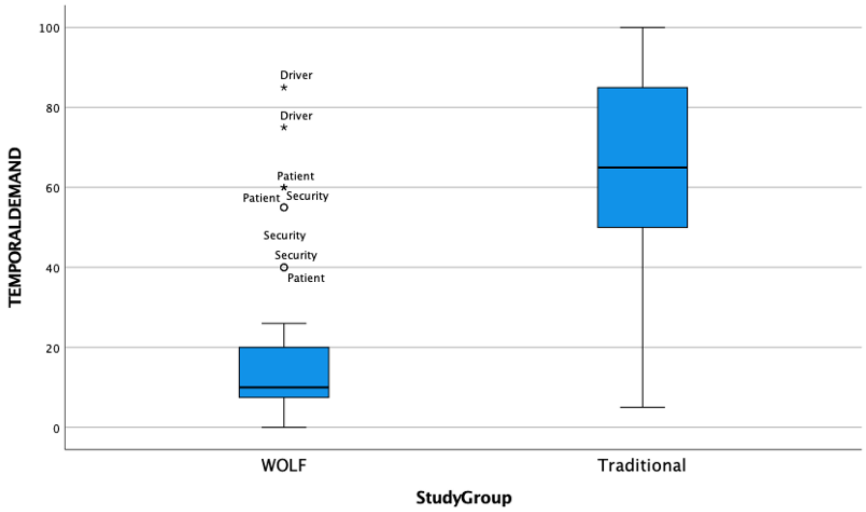


Figure 6: The time it took for patients to be evacuated off the battlefield was significantly shorter for the WOLF group, but it FELT longer to some of those group members, according to our findings. This indicates that some functions of interaction with the Hunter WOLF can be improved and made smoother to improve the overall experience.

cles like the WOLF so that they “fit” our customers well. Improving the loading and unloading of patients, for example, might very well improve the current findings of high temporal demand in addition to the physical demand experienced by our participants.

OVERALL PERFORMANCE

At the end of each exercise, we asked each individual how they thought they did overall in performing their mission of evacuating patients safely. Overall performance, as measured by each participant’s self-assessment using NASA TLX, differed significantly between the two groups, t(114) = 10.17, p < .001, d = 23.49. Based on observations and discussions with training staff, the groups who evacuated patients using the WOLF were more likely to reach their objectives with less safety issues, and had greater overall communication and teamwork than the groups who evacuated patients using traditional litters. We infer from these anecdotal observations that the WOLF afforded teams greater freedom of movement with less overall effort, which resulted in better overall performance.

We can also consider patient survivability as a measure of overall performance. When Marines and corpsman failed to effectively treat their wounded patients, or when they took too long to administer aid or move their patients to safety, the Coyotes would mark the mock patient as deceased. Three WOLF patients were designated as deceased, or around 11% of the total of 28 patients of the WOLF evolutions, whereas two patients during the traditional litter carrying evolutions were designed as deceased, which was 25% of the 8 patients involved in the traditional litter evolutions. It is nearly impossible to directly correlate the mock patient outcomes to the presence or absence of the WOLF. The factors we have measured above (i.e., physical workload, mental workload, effort, frustration, and time), however, do allow us to infer some relationship between the manner in which patients were evacuated during our exercise, and whether or not they survived the evolution.

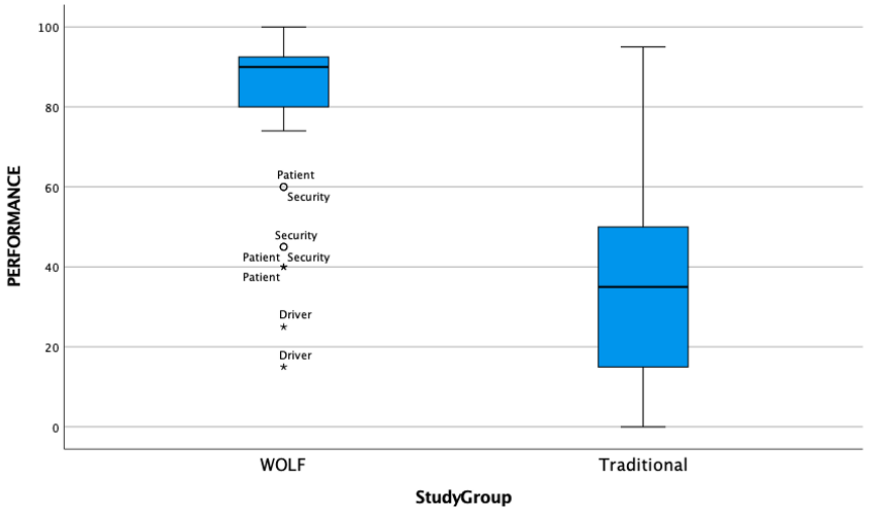


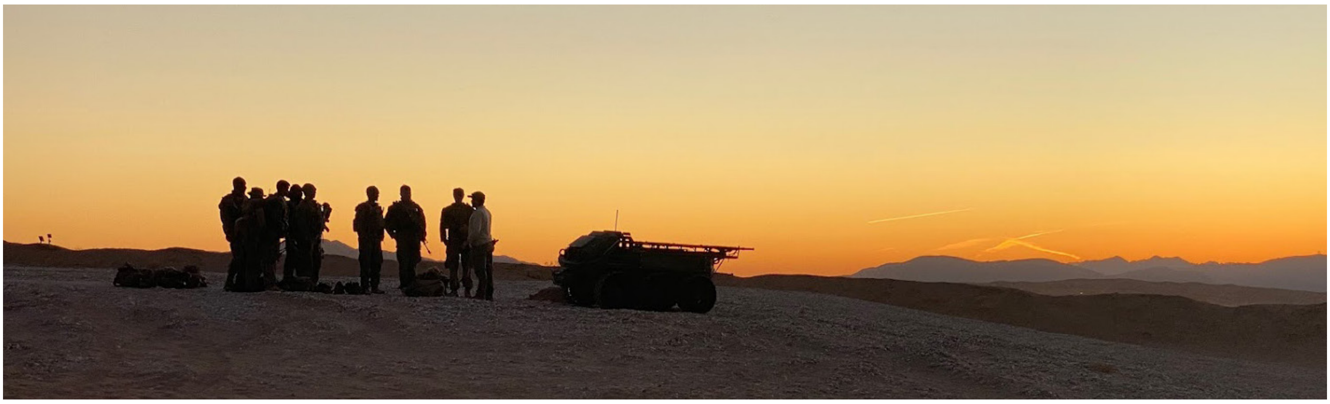
Figure 7: Overall performance, as measured by each group’s ability to evacuate their patients to the designated location.

CONCLUSION

The purpose of 6.5 activities on the RDT&E spectrum is to demonstrate the possible in order to inspire what can be done, and refine designs so that they best meet the needs of the operational customer. Evaluations are guided by questions to be answered rather than research hypotheses. The methods we employed for this evaluation of the autonomous casualty evacuation concept, therefore, reflect our interest in understanding where benefits can be attained with vehicles like the WOLF in dynamic operational settings, and likewise where situations favor other technologies or analogs.

Our evaluation of the WOLF for casualty evacuation demonstrated several benefits over traditional methods. Future conflicts may feature situations in which capabilities afforded by autonomous vehicles may be a significant factor in determining successful outcomes. As is common in robotic and autonomous systems, however, the manner in which these systems are designed and employed can play a large role in their effectiveness, and this evaluation demonstrated several examples of this as well. Getting the technology right is only half the battle.

Ultimately, how the US Navy and Marine Corps plan to manage casualty evacuation in future conflicts is partially informed by results from studies such as this. Taking science and technology out of the constraints of laboratory environments and into environments that mimic real life is critical to the success of fielding innovative solutions. This is the purpose of 6.5, system development and demonstration activities on the RDT&E spectrum.



An F-35 Joint Strike Fighter completes a test mission at Pax River Integrated Test Force. The F-35C is one of several high-performance aircraft that can make pilots susceptible to the effects of hypoxia. Training to recognize and mitigate these effects is central to maintaining and improving naval aviation safety. (Photo courtesy of Lockheed Martin)

TASKING IN THE CLOUDS

Rethinking aircrew tasking during normobaric hypoxia training

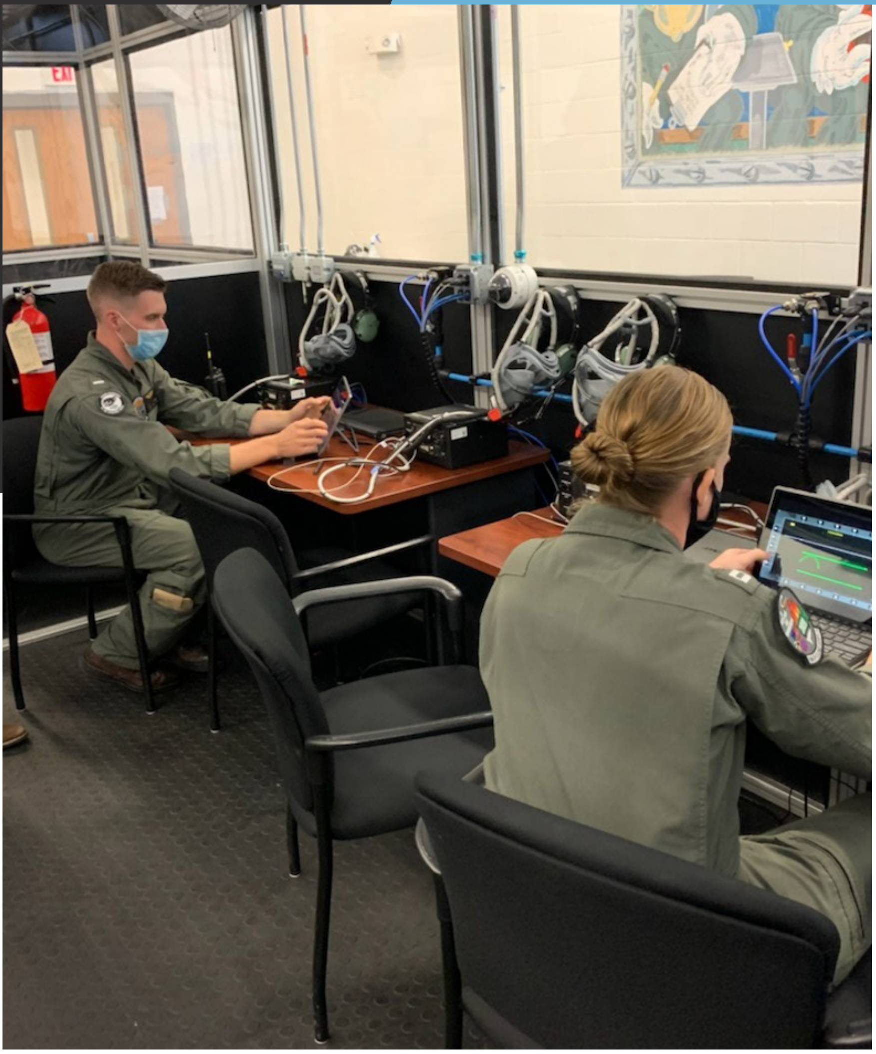
By: Mitchell J. Tindall, Beth F. Wheeler Atkinson,
LCDR Brennan D. Cox, LCDR Lee W. Sciarini, & LCDR Daniel L. Logsdon

In 2016, the Navy opted to decommission a piece of training history. Due to health and safety concerns, as well as lifecycle maintenance impacts resulting from obsolescence, the Low Pressure Chambers (LPCs) ceased to be a part of Navy aviation survival training. In conjunction with this shift, the Chief of Naval Operations directed the Naval Aviation Survival Training Program (via CNAF M3710.7) to adjust training requirements to use the Normobaric Hypoxia Trainer (NHT) in lieu of LPCs. The NHT allows the Aviation Survival Training Centers to train multiple aviation personnel in hypoxia recognition and emergency procedure familiarization, for initial and aircraft class 2 and 4 aircrew refresher training (i.e., individuals who fly in pressurized aircraft that only use oxygen masks during rapid decompression or other emergencies). The NHT was designed to provide mask-off aircrew a hypoxia training solution that did not employ pressurization and therefore avoided significant risks to safety, while still allowing aircrew to experience how altitude impacts physiology and breathing. While the Naval Aviation Survival Training Program employs lectures and practical application training across major curriculum topics, hypoxia exposure training for aircrew have traditionally lacked scenario-based simulation to provide a medium for a comprehensive training experience.

As a part of a multi-year collaboration across several Navy organizations, a software application was designed. The objective was to increase training fidelity and trainee workload by providing students with aviation-relevant tasking to complete while experiencing the effects of hypoxia to illustrate cognitive deficiencies and physiological symptoms. The collaborative team included personnel from the Naval Survival Training Institute, Aviation Survival Training Centers, the Naval Air Warfare Center Training Systems Division, Naval Postgraduate School, and industry. Through this collaboration, a team of researchers including both Aerospace Experimental Psychologists and Aerospace and Operational Physiologists leveraged expertise in training, human factors, and human physiology to iteratively design and develop an innovative capability to enhance aircrew training using engagement with aviation relevant tasking to illustrate the dangerous impacts of hypoxia.

Replacing the Chamber: The Normobaric Hypoxia Trainer

The NHT safely exposes aircrew and aviators to a simulated high-altitude environment by lowering oxygen and not manipulating pressure during a training session. Oxygen levels are adjusted from an equivalent altitude of 12,500 feet mean sea level and then slowly increasing to 25,000 feet over approximately 25 minutes. This exposure to an



The final software is tested at the Aviation Survival Testing Center onboard Naval Air Station Pensacola, Florida. Naval Opertional and Aerospace Physiologist interns LTJG Derik Kincaid and LT Alicia Jordan serve as test subjects of the software.

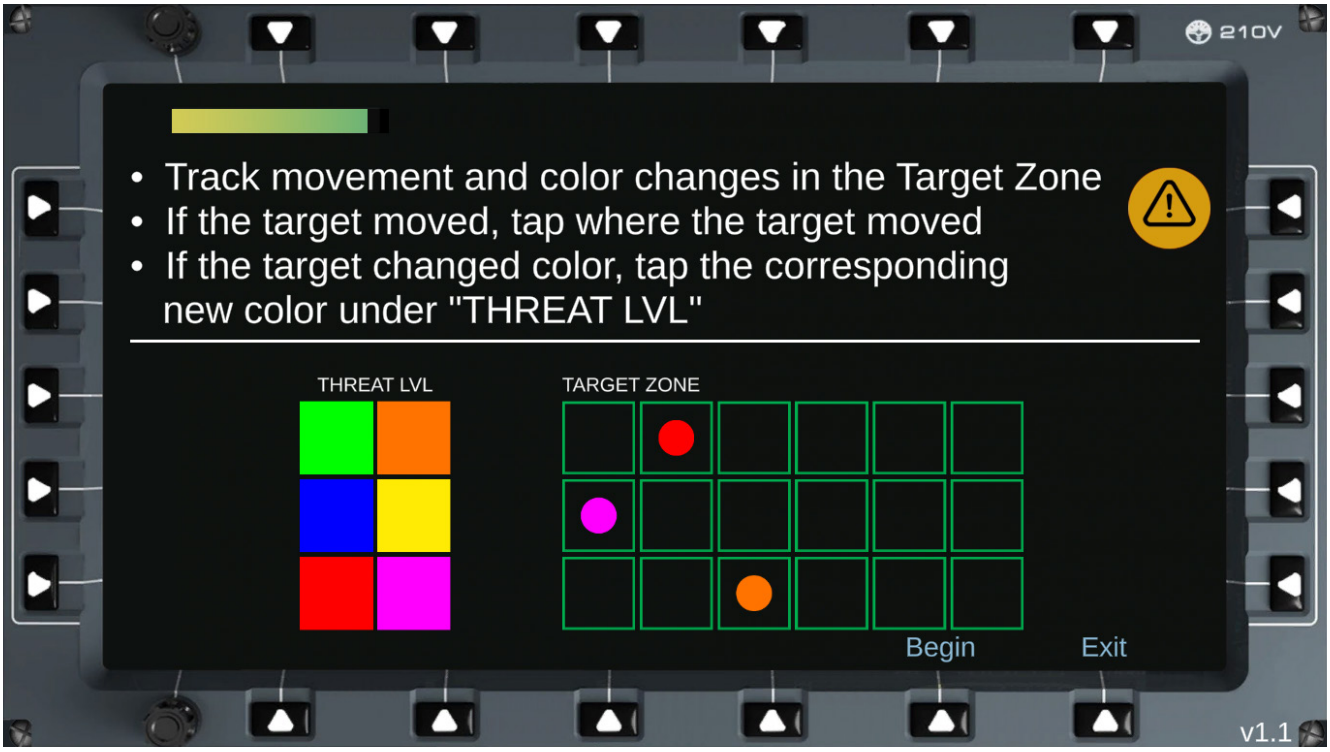
oxygen-deprived environment is one of the dynamic altitude threats faced by aviators and can result in rapid loss of mental, physical, and/or psychomotor abilities. Given the catastrophic consequences of hypoxia in the aviation domain, and the lack of automated systems for intervention, this experience offers a critical opportunity to increase the likelihood of recognition of symptoms and execution of emergency procedures that mitigate the potential for loss of life and aircraft. The NHT delivers advancements beyond LPC legacy training in both safety and training fidelity. While implementing a flight station simulator was an intuitive decision to increase training fidelity for pilots, developing operationally relevant tasking that engages initial and refresher aircrew students required extensive requirements gathering, analysis, and creative design.

The original NHT design requires trainees to sit at an aircrew station allowing them to complete operationally representative tasks on computer tablets. As symptoms of hypoxia are experienced, trainees activate recovery air to simulate emergency procedures with a goal of: 1) learning to recognize hypoxia symptoms, 2) learning proper procedures to alleviate hypoxia symptoms, and 3) learning to efficiently perform emergency procedures using actual aircraft life support equipment to prevent hypoxia-related mishaps. During the procurement of the NHT, each device was delivered with tablet computers for training aircrew positions; however, the tablets did not include simulation software to offer domain-relevant tasking. This gap resulted in minimal cognitive or psychomotor tasking for aircrew, and was identified by the NASTP Trainer Management Team (TMT) as a strategic priority. Executing an operationally relevant task provides students with a demonstration of the impacts of this altered physiological state, which is essential to ensuring they can recognize symptoms and execute emergency procedures.

Refining Concepts for Aircrew Tasking

The Hypoxia Awareness Trainer (HAT) application offers a cycle of distractor tasks that include basic Stroop, double Stroop, advanced Stroop (i.e., multitasking Stroop), and target tracking tasks that are similar to challenges Naval aviation indoctrination trainees are exposed to at the schoolhouse. Stroop tests were initially developed as tools for researchers to ascertain cognitive effects of a manipulation on a subject in psychological research. They do this by requiring subjects to process multiple pieces of information simultaneously, a core skill set necessary for all military aviators. For this reason, the Stroop test provided non-domain relevant distraction tasking that could tap important foundational skillsets of Navy aviators during hypoxia training. In addition to providing tasking that is familiar to aviators, the software application is designed to have a similar look and feel to an operator's Heads Down Display. By replicating this onboard hardware and software system, the interaction requirements are highly intuitive. As students interact with the application, HAT offers an opportunity to ensure that students remain engaged and focused on relevant tasking that might distract them from identifying their symptoms. Training is considered successful if students, while immersed in tasking, recognize their symptoms and engage their emergency procedures. The primary benefit of engaging with non-domain relevant tasks that might distract from symptom identification is the similarity of the cognitive workload and tactical engagement required in operational settings; this is a crucial component of robust and effective aviation safety training.

This iterative design and development program was initiated as part of a Naval Research Program funded Human Systems Integration / Modeling and Virtual Environments project at the NPS. The Naval Research Program is categorically



- Track movement and color changes in the Target Zone
- If the target moved, tap where the target moved
- If the target changed color, tap the corresponding new color under "THREAT LVL"

This task requires operators to monitor multiple target positions and colors while under the controlled effects of hypoxia. These aparatuses help aviators recognize the early effects of hypoxia and train to mitigate its effects.

programmed for studies and analyses in support of Research, Development, Test and Evaluation Management Support (Budget Activity 6) and serves as a focal point, stimulus, and major source of strategic, tactical and operational thought within the Navy communities while supporting students' capstone and thesis projects as well as faculty projects.

During this phase of the effort, NPS conducted a survey to identify operationally relevant aircrew distractor tasking across multiple platform types. Their results informed development of a prototype application that underwent testing and evaluation by end users. As a follow on effort, a Defense Health Program (DHP)-funded effort conducted workshops to finalize the application design grounded in the original task concepts but generalized across platforms. As a Budget Activity 7 program, the DHP enabled the development efforts to upgrade the NHT which, at the time, was in the process of being delivered to each of the eight ASTCs.

Based on the initial NRP work from NPS, the approach was to create platform agnostic tasking as opposed to platform specific to ensure the wide variety of students – from indoctrination students to refresher trainees across platform types – could perform tasks and minimize lifecycle sustainment costs. Employing a hybrid heuristic-survey-based usability evaluation process and agile software development allowed for rapid prototyping and increased user feedback.

This effort concluded with a final software development build that addressed all feedback provided. The software is awaiting final approval for use by the NASTP TMT and Integrated Product Team (IPT) with a focus on ensuring cybersecurity adherence, prior to installation across all eight ASTC NHTs. In addition to the strong positive response by Navy stakeholders, a demonstration for U.S. Air Force personnel resulted in early interest regarding how the technology could support safety training in their environments.

The HAT solution provides other advantages for both research and future training. First, the device actively captures student performance at baseline, throughout the training profile, and at recovery. One area for further research is broadening the understanding of both patterns of individual differences and the stability or changes in symptom profiles over time. While current training operating procedures do not capture data for longitudinal analyses, the HAT software was designed to support this data capture when desired and appropriate. Second, as alluded to previously, the HAT application improves engagement and increases training fidelity for students. In legacy training solutions, students engaged in "Pensacola patty-cake" or interacted with child shape puzzles in attempts to demonstrate cognitive degradation. Similarly, with the delivery of tablet devices without appropriate software, solitaire provided the only tasking opportunity for aircrew distraction; these prior approaches to providing student tasking resulted in limited engagement and poor training fidelity.

A MH-60S Seahawk from the “Golden Falcons” of Helicopter Sea Combat Squadron (HSC) 12 deploys flares during an air power demonstration as part of the U.S. Navy’s forward-deployed aircraft carrier USS George Washington’s (CVN 73) four-day tiger cruise. Safety in Naval Aviation remains a critical issue and is the principal focus of the Naval Safety Command. (U.S. Navy photo by Mass Communication Specialist Seaman Apprentice Oscar Albert Moreno Jr./Released)

LEARNING SAFETY

How an organization like Naval Aviation learns to fly safely.

By: CDR Jefferson D. Grubb
Human Factors Branch Head, U.S. Naval Safety Command

Although much of the research in which Navy Aerospace Experimental Psychologists (AEPs) directly participate is funded with Research, Development, Test and Evaluation (RDT&E) appropriations, AEPs also perform research as part of their duties at non-acquisition commands. The following is an example of one such project.

The Naval Aviation Ski Slope Chart

The newly christened U.S. Naval Safety Command (NSC) is the latest iteration of an entity the Navy created in 1951 as the Naval Aviation Safety Activity. As part of its mission to preserve warfighting capability, NSC collects and analyzes mishap reports. One of its most famous information products is a chart that plots the annual rate of major Naval Aviation mishaps since 1950 (see Figure 1). This chart, informally known as the “ski slope,” shows a steep initial decline that gradually tapers to a relatively low, steady rate by the 1990s. Presenters typically annotate this curve to show the timing of significant safety developments.

Although the ski slope is ubiquitous in safety presentations, its message is rather ambiguous. Clearly, the current mishap rate is much lower than it was in the 1950s. However, most of the drop in the mishap rate occurred before 1965. When NSC personnel ask groups of prospective commanding officers (PCOs) why the recent rate looks so steady, the PCOs frequently debate whether it shows a decay in our safety culture or that the mishap rate has reached some lower limit. When NSC personnel subsequently ask what annotated events drove down the mishap rate, most of the PCOs just smirk. Although every presenter’s version of the ski slope shows annotated events, the specific events vary widely between presentations. Moreover, the timing of commonly listed events also varies widely. Regardless, the annotations rarely line up with obvious changes in the mishap rate curve. In the end, ski slope charts show that we did stuff and that we got better. However, it isn’t clear whether the two are related and, disturbingly, we don’t seem to be getting much better anymore.

Measuring of the Slope

To better understand the ski slope curve, the author obtained the raw annual major mishap rate data from NSC’s Knowledge Management and Safety Promotions Directorate. Although ski slope charts usually show mishap rates since 1950, the full data set goes back to 1922. As Figure 2 illustrates, the earlier data appears nosier, but the overall shape of the full plot looks remarkably similar to the standard post-1950 plot. In this context, both the steep decline in the 1950s and the

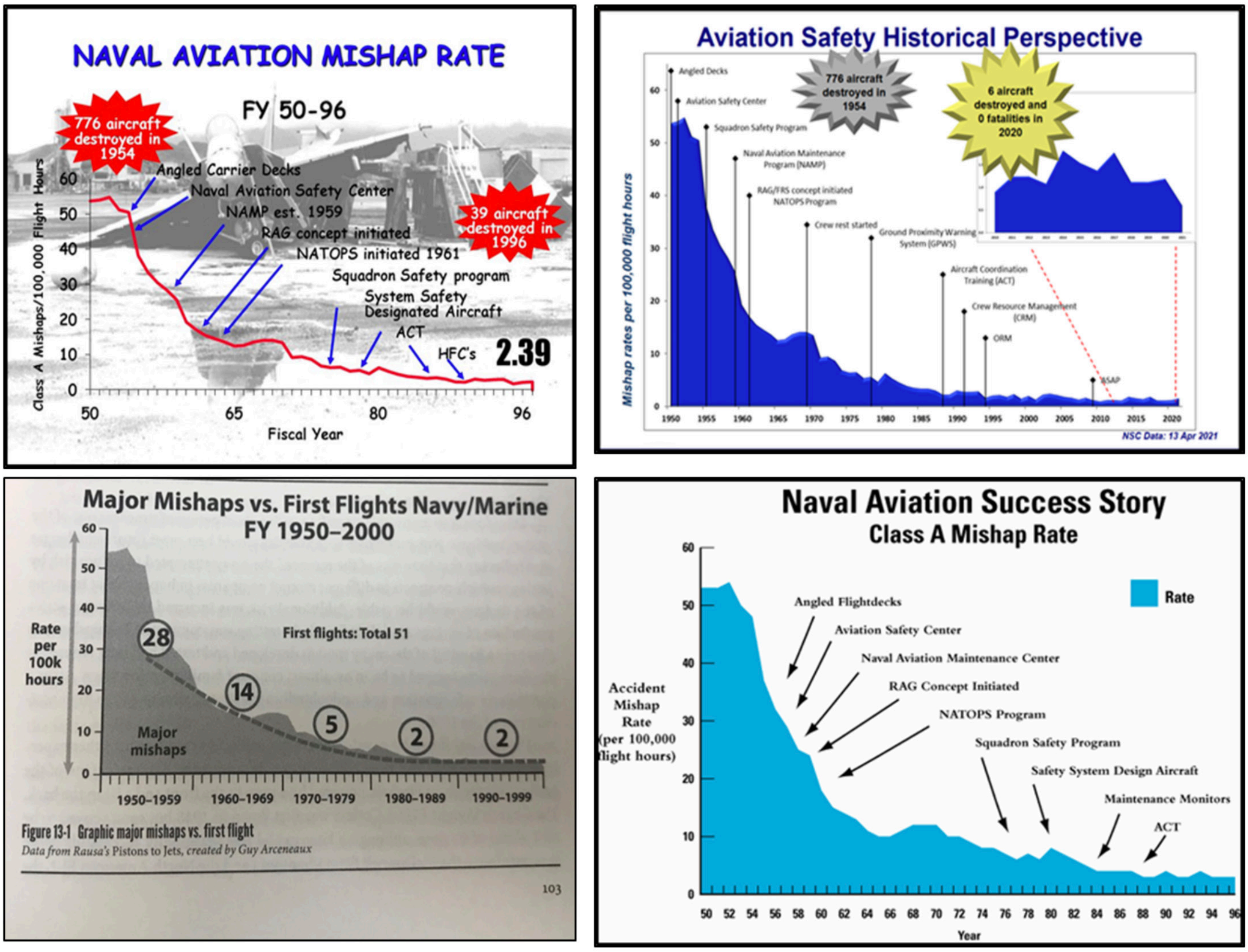


Figure 1: Four examples of the "Ski Slope" chart. Although the curve is the same in each chart, the annotated events, and even the indicated timings of commonly listed events, vary between charts

recent steady trend appear to be part of a consistent, general pattern that goes back to the earliest days of Naval Aviation. To characterize this pattern, the author fit linear, logarithmic, power law, and exponential models to the full data set. The exponential model provided the best fit, explaining 97% of the variance in annual major mishap rates since 1922. Importantly, this indicates that although the mishap rate has been declining by smaller absolute amounts over time, it has continued to decline by a remarkably constant relative rate. As Figure 3 shows, although this model fit the data remarkably well overall, there were extended periods in which the actual mishap rate consistently fell above or below the model's predictions.

To determine whether any of these periods represented significant changes in the overall pattern of mishap rates, the author detrended the mishap rate data and then conducted a change point analysis using the pruned exact linear time (PELT) algorithm for mean and variance (Killick & Eckley, 2014). These change points are plotted on the raw and detrended time series in Figure 4.

The Ski Slope is a Learning Curve

The main finding of this analysis was that major aviation mishap rates follow a negative exponential curve. Such curves are characteristic of eliminative learning curves (Heathcoat, et al., 2000). This strong (R2 = 0.97) trend suggests that ***organizational learning is the principle driver of improvements in Naval Aviation safety***. That is, Naval Aviation safety has historically improved through the gradual development, promulgation, adoption, and retention of techniques, procedures, policies, and material designs that enable aviation personnel to operate more safely. Individual technical or cultural interventions, or so-called "silver bullet" solutions, have rarely, if ever, led to discrete improvements in safety.

Importantly, saying that Naval Aviation demonstrated organizational learning is different from saying that Naval Aviation is a "learning organization." Organizational learning is a process whereby an organization's behavior changes with experience. While such change can be intentional and beneficial to the organization, it can also be haphazard or even maladaptive. In contrast, a learning organization refers to an organizational structure designed to promote active learning of behaviors that advance the organization's mission.

Although the business literature tends to present learning organization status as binary, the change point analysis suggests that Naval Aviation safety fluctuates with its degree of adherence to learning best practices. The analysis presented here identified two consecutive epochs of unusually steep decline in the mishap rate beginning in 1954 and continuing through 1966. This period corresponds to the release of

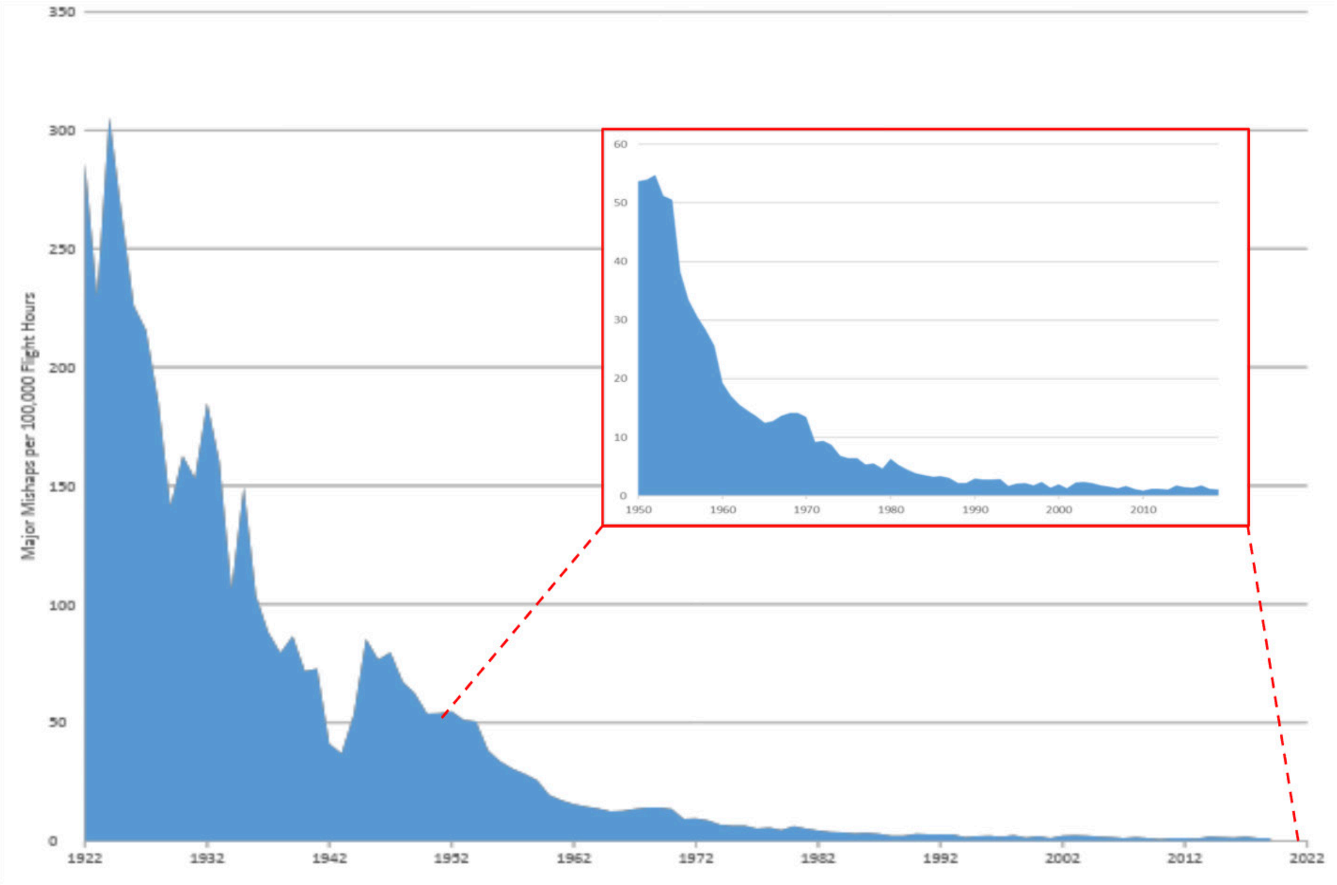


Figure 2: Naval Aviation annual major mishap rate per 100K Flight Hours 1922 - 2019. For overall comparison, the inset shows the post-1950 data, which are what typical ski slope charts present.

the Flatley Report and the subsequent holistic reformation of Naval Aviation training and organization (Dunn, 2017). Among other things, this period saw the standup of the Naval Aviation Training, Operations, and Procedures Standardization (NATOPS) program, the Naval Aviation Maintenance Program (NAMP), and Fleet Replacement Squadron (FRS) concept. These and other efforts during this period promoted the collection, dissemination, and enforcement of best practices across the Fleet.

In contrast, the change point analysis presented here showed five epochs in which the mishap rate increased. These correlate with periods in which adherence to best practices was likely challenging due to sustained combat operations or acute budgetary turmoil, where operational demands unexpectedly outstripped resourcing. Naval leaders must make short-term trades to cover the demands, often at the expense of practices that ensure the long-term sustainability of the force.

Likewise, the rapid, large-scale introduction of disruptive technology (e.g. jet aircraft, large deck aircraft carriers, etc.) likely forced Fleet personnel to learn important characteristics of these technologies from scratch. No one can follow best practices in these circumstances because no one has yet learned what they are.

In summary, this analysis presented here indicates that leaders who want to improve safety should approach the problem as a learning management exercise. Development of new safety solutions is important, but the reduction in mishap rates ultimately depends on their dissemination and retention. Structuring the organization to promote learning may even allow the Navy to better capitalize on solutions it has already developed at great cost. Ultimately, the Fleet cannot benefit from a lesson learned unless it truly is learned.

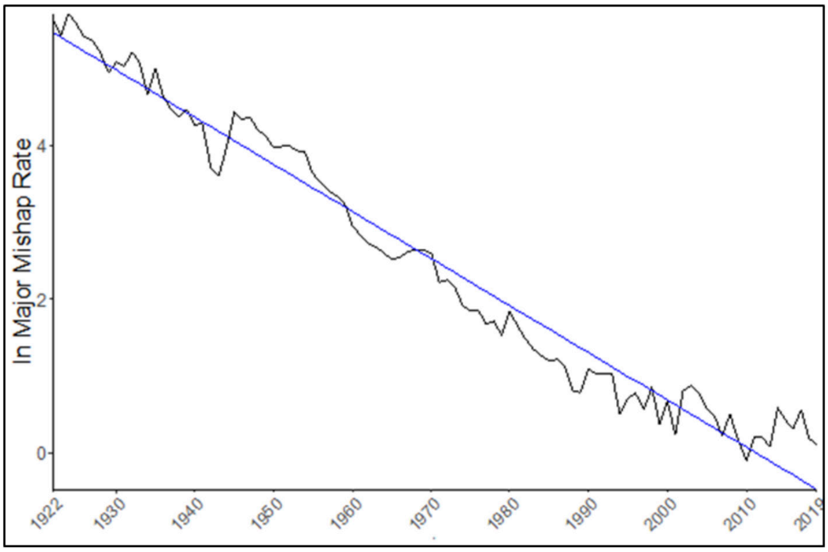


Figure 3: Semi-Log Plot of Naval Aviation Annual Major Mishap Rates. Blue line indicates best-fit exponential model.

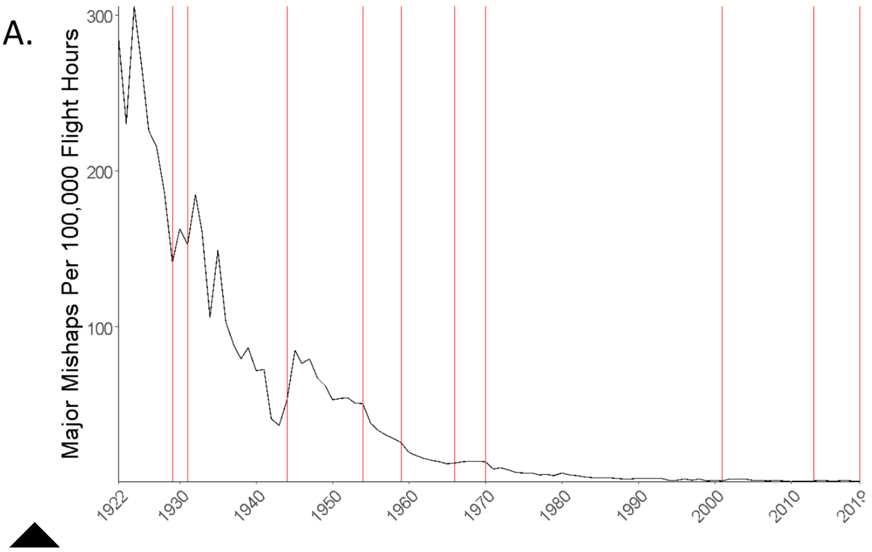


Figure 4: Change Points of Naval Aviation Annual Major Mishap Rate. Red lines indicate times at which the mean and variance of the deviation of the mishap rate from the best-fit exponential model changed significantly.

REFERENCES

Dunn, R. F. (2017). Gear up, mishaps down: the evolution of naval aviation safety, 1950-2000. Naval Institute Press.

Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. Psychonomic bulletin & review, 7(2), 185-207.

Killick, R., & Eckley, I. (2014). changepoint: An R package for changepoint analysis. Journal of statistical software, 58(3), 1-19.

The author wishes to thank LCDR Brennen Cox for his comments and recommendations on the interpretation of the analyses presented here.

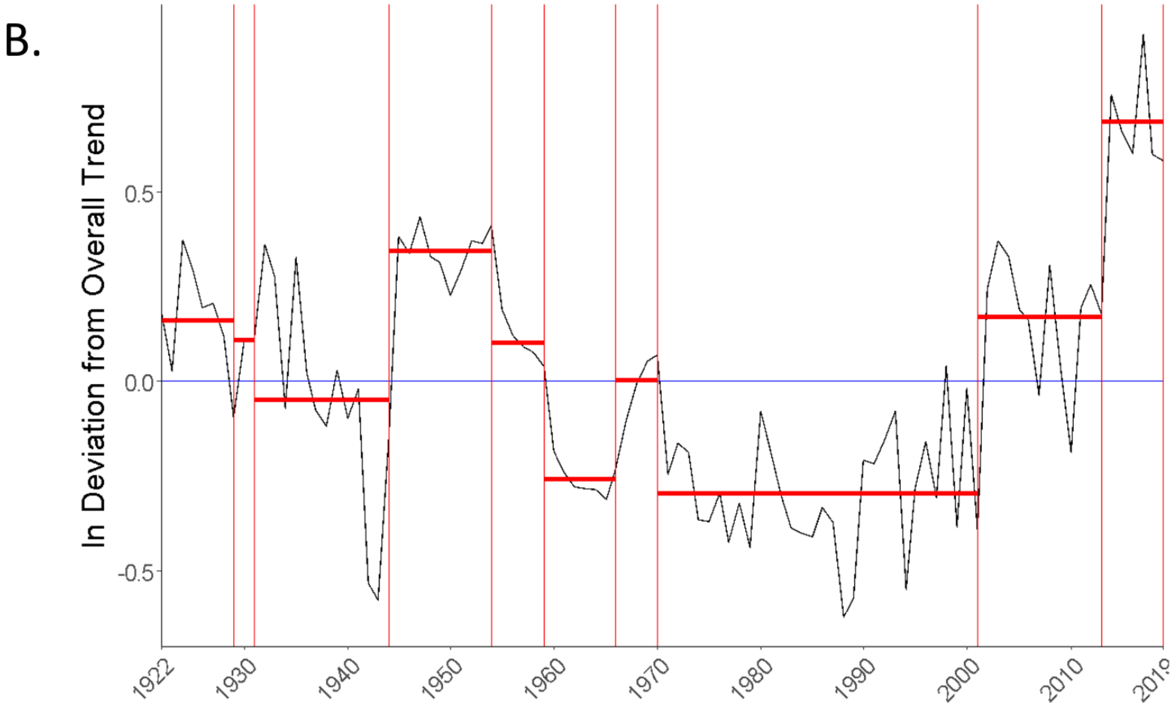


Figure 5: Change Points Plotted on Detrended Data. This graph shows the results of the same analysis as that in Figure 4, with change points plotted against detrended data to highlight the direction and magnitude of the change relative to the best-fit exponential model.

MEET AN AEP

LT Sarah Beadle, AEP #164, discusses why they chose a career in the US Navy as an Aerospace Experimental Psychologist

The decision to leave civilian life and join the military is one that involves a lot of personal choice and preferences. There is no "standard" servicemember, and there is certainly no standard uniformed scientist, despite what some movies or books may depict.

In this series, we spotlight individual AEPs to learn more about them in a one-on-one interview format in order to narrow that gap, and foster relationships and collaboration across our community. In this issue we will meet LT Sarah Beadle. Sarah recently graduated from Naval Aviation Flight Training and is preparing to report to her first duty station at the Naval Air Warfare Center, Aircraft Division at Naval Air Station Patuxent River, MD as an Aerospace Experimental Psychologist.

What is your academic background?

I went to Simpson College in Indianola, Iowa for my Bachelor's degree in Psychology and Neuroscience. My first exposure to research was good timing- Simpson had purchased an eye tracking system the year I started and I was funded to set it up and teach people how to use it. I attended Clemson University for my Master's and Doctoral degrees in Human Factors Psychology. My industry experience comes from an Oak Ridge Institute for Science Education traineeship with the U.S. Army Aeromedical Research Lab and internship at Lockheed Martin Rotary and Mission Systems.

My graduate training and expertise is in simulator sickness in head-mounted displays. Specifically, I studied how latency impacts symptom reports and adaptation to simulator sickness. I had other exposure to work on wearable devices and spatial disorientation projects. Lucky timing, while at Clemson I also got to see two national championship football teams!

What made you interested in pursuing a Doctoral degree in human factors?

I chose a degree in human factors for its inherently applied nature. When I was in college, I got the liberal arts training that every problem is interdisciplinary in nature and it benefits to be on a collaborative team. I chose Clemson as a program because it was required that I take classes outside of psychology- in fields like industrial engineering and human-centered computing, to be well-rounded. To me, human factors is psychology at it's most challenging- how people interact with new technology and how fields like aviation employ psychology.

LT Sarah Beadle receives her "wings of gold" at a ceremony on February 11th, 2022.

How did you learn about the AEPs?

I have the privilege of being an AEPs who got to be trained by an AEP. At Clemson, I worked under Dr. Eric Muth (AEP #109). Eric spent a lot of time talking about how the Navy shaped his career and approach to science. Due to the nature of our research, I read the work of many AEPs, and was particularly molded by the work of CDR Bob Kennedy (AEP #10) studying simulator sickness and motion sickness.

As I progressed in my graduate degree, I met more and more active duty and civilian Department of Defense scientists and started to see why they chose to combine science and military service. As I was finishing up my degree, I was funded on a Small Business Innovation Research project with the Naval Survival Training Institute examining flight simulation as a way to modernize teaching about spatial disorientation and mishaps. That project taught me a lot about government research, but mostly how foreign I was to aviation. I realized that if I wanted to keep in this line of work, the exposure to aviation was something I needed to pursue.

What was the most challenging point of AEP training?

I found that I had to get comfortable with being uncomfortable during water survival training! The training is physically and mentally tough, and was my first big

LT Sarah Beadle prepares for a flight in the T6 Texan II, the US Navy's fixed wing training aircraft

LT Sarah Beadle receives her "wings of gold" and officially becomes AEP #164. She is accompanied by her academic advisor, Dr. Eric Muth, who retired from the US Navy as AEP #109.

exposure to high-risk training. I really didn't like staying underwater or treading water with gear on, so I had to get over that before I got to the infamous Helo Dunker. The confidence gained in the pre-flight phases of flight school certainly helped me keep a level head in the months that followed.

What was your most memorable moment during training?

I loved doing aerobatics in the T-6A. I was so nervous my first flight because I had never flown in anything other than a commercial aircraft and then was suddenly in the front seat of an ejection seat turboprop. Experiencing G-forces and being upside down in an aircraft for the first time is unforgettable. Luckily, the fears I had about experiencing airsickness based on my research background didn't come into play!

Where do you see yourself in 10 years?

I really enjoying evaluating training technologies and learning how we get feedback from the fleet to push future needs. In 10 years, I hope to see myself in a leadership role working closely with fleet aviators, flight officers, and enlisted leaders to push science and technology development that supports their needs. It has been really exciting to see how virtual reality and HMDs are being employed by the Navy right now and I'm interested in shaping that continued effort in the future. Regardless of what billet I occupy at that time, I see myself staying hungry to travel to the next challenging environment- hopefully spending time overseas, on a ship, and chasing more operational experiences.

LT Sarah Beadle receives her graduation certificate for Naval Introductory Flight Evaluation from Captain Moreno, Commanding Officer of Naval Aviation Schools Command, Pensacola, Florida.

USNAEPS AWARDS

Our best and brightest are honored with these annual awards

CDR (Ret.) Michael Reddix (AEP #100) won the Paul R. Chatelier Lifetime Achievement Award for his sustained efforts over a 22-year active duty career, CDR Reddix forged a path of sustained scientific, managerial, and leadership excellence. His substantial research contributions in the vision sciences have had lasting impact, including a revision to naval aviation's color vision screening standards as well as a requirements-to-acquisition laser eye protection solution for the US Coast Guard. As Officer-in-Charge, he was directly responsible for managing the Base Realignment and Closure of the Directed Energy Bioeffects Laboratory, culminating in 2010 with the establishment Naval Medical Research Unit – San Antonio. He later served as Executive Officer of Naval Medical Research Unit – Dayton, where he maintains a post-military career as senior technical advisor.

LCDR Stephen Eggan (AEP# 143) won the Michael G. Lilienthal Leadership Award for his exemplary leadership and dedication to supporting the health and readiness of warfighters across the CENTCOM, AFRICOM, and EUCOM areas of responsibility (AORs) during the COVID-19 pandemic. He stood up the only forward deployed COVD-19 operational risk reduction program supporting more than 10,000 active duty, U.S. Embassy, and partner nation personnel across 12 countries in the AFRICOM and CENTCOM AORs providing rapid detection and response capabilities to outbreaks and preserving force health protection and operational readiness. Additionally, in support of Operation Allies Refuge, he coordinated deployment of 30 medical teams to provide 24/7 COVID testing to more than 4,300 Afghan refugees.

LT Aditya Prasad (AEP #156) won the Robert S. Kennedy Award for Excellence in Aviation Research for his efforts over his time at Naval Air Systems Command 4.6.5, NAS Patuxent River, MD to research the factors involved in Air-Ground Mishaps and how they impact fleet readiness. He presented his findings at the 2019 Military Health System Research Symposium and the 2020 United States Navy Aeromedical Conference, and he was invited to give an oral presentation of his findings to the 2020 Military Health System Research Symposium.

LT Todd Seech (AEP# 153) won the Robert S. Kennedy Award for Excellence in Aviation Research for his pioneering efforts in advancing Naval aviation training instructional design and evaluation. LT Seech's research efforts have transformed key areas of the naval aviation training pipeline and provided critical insights into how artificial intelligence (AI) and virtual reality technologies can be most effectively integrated into Navy and Air Force flight training to reduce live flight time training by an estimated 30%, significantly reducing overall time to train and producing cost savings to the Department of Defense of $328 million annually.

CDR (Ret.) Henry Phillips (AEP#119) won the Paul R. Chatelier Lifetime Achievement Award for significant and enduring contributions in aviation selection, training, and combat casualty care. He revolutionized aviation candidate selection by transforming the Aviation Selection Test Battery (ASTB) to computer adaptive testing, shortening testing time by 50% while increasing measurement precision and predictive validity, saving the Naval Aviation Enterprise more than $52 million annually. He directed $67 million in scientific programs to deliver next generation casualty care and resilience training tools to the Marine Corps and oversaw delivery of a suite of training tools to address critical gaps in Sailor readiness in response to the USS McCain and USS Fitzgerald mishaps.

AEROSPACE EXPERIMENTAL PSYCHOLOGY

WHERE SCIENCE AND SERVICE UNITE

TAKE YOUR PHD TO NEW HEIGHTS IN THE US NAVY AS AN AEROSPACE EXPERIMENTAL PSYCHOLOGIST!

1- HUMAN FACTORS

We conduct research and analyses of naval aviation platforms and systems, using cutting edge human performance techniques.

2- TRAINING & SIMULATION

We evaluate and develop training technologies for all elements of aviation platforms that feature state-of-the-art in virtual and augmented reality.

3- PERSONNEL SELECTION

We oversee the selection test battery used for identifying ideal candidates for naval aviation by predicting success in the training pipeline.

US naval aviation is a fast-paced and dynamic domain in which you can immediately apply your skills and education while serving your country.

The opportunities and benefits- from generous pay and allowances to once-in-a-lifetime experiences- are abundant.

Come see how you can make a difference! Contact me to learn about our recruitment process at:

www.navyaep.com

Captain Joseph Cohn, PhD
AEP Specialty Leader

US NAVAL AEROSPACE EXPERIMENTAL PSYCHOLOGY SOCIETY | MAIL: INFO@NAVYAEP.COM | HTTPS://WWW.NAVYAEP.COM