

Bridging the Gap: The Critical Need for Crew Resource Management (CRM) Training in Space Domain Operations

Jason Goldberg

Riverside Research Institute

Fairfax, VA, USA, JGoldberg@RiversideResearch.org

Timothy Waltz

Riverside Research Institute

Fairfax, VA, USA, TWaltz@RiversideResearch.org

ABSTRACT

Human error is ubiquitous and inevitable. As automation and artificial intelligence tools become more prevalent, the role of human operators becomes more critical in managing complex systems during normal and off-nominal situations. Addressing human factors is essential to ensure safety by trapping human error when it occurs. Minimizing the consequences of human error is crucial for maintaining operational effectiveness and optimizing performance in increasingly complex environments. Furthermore, it helps preserve strategic assets by protecting valuable space infrastructure. As the operational environment of space becomes increasingly complex and technology-dependent, the role of human operators remains pivotal in ensuring mission success and infrastructure safety.

To ensure the success of space operations in increasingly complex environments we introduce the applicability of Crew Resource Management (CRM) principles, rooted in behavioral psychology. This includes activities ranging from ground-based sensor management, centralized scheduling and tasking, coordination efforts, to spacecraft control and satellite operations. CRM is the process used by crew members to identify existing and potential threats to safe operations and to develop, communicate, and implement plans and actions to avoid or mitigate perceived threats. Despite the demonstrated benefits of CRM in aviation and manned spaceflight, it has yet to be systematically adopted across satellite operations and space domain awareness (SDA) functions. We introduce a pathway to enhance the human operational layer within space domain activities by integrating CRM tenets, thereby improving team performance, safety, and mission success. This is accomplished by providing training that focuses on group dynamics including assertiveness, leadership, mission analysis, situational awareness, interpersonal communications, adaptability/flexibility, and decision-making. Functional SDA or satellite operations training have traditionally focused on the human-machine interface, emphasizing technological tools and user preferences. While there is substantial scientific data on the technologies and tools used by operators, there is a notable gap in applying behavioral psychology principles to improve team dynamics and human performance in space operations. We submit four case studies illustrating the applicability of these principles to space operations. Integrating these principles into training at all levels is a vital piece to improving upon those team dynamics.

This paper argues for the urgent implementation of CRM training for space domain operations personnel by synthesizing findings from historical case studies, behavioral psychology principles, and human factors policy frameworks. Drawing on FAA guidance, human error research, and CRM case studies, we present a comprehensive paradigm and actionable pathway for integrating CRM into space operations training programs.

1. Introduction: Background and History of CRM

Dating to the dawn of *Homo erectus*, the pursuit of improving the human condition is perpetual. Distinctly human contributions to this pursuit in all its forms provide a backdrop to our topic. Throughout history humans have been both the “how” and the “why” behind evolutionary spirals of development, exploration, and other achievements that advance the human condition. With the advent of technology, humans were able to differentiate themselves by creating systems that multiplied productivity of individuals and groups. When technological spirals began, humans became part of an ecosystem of systems with sentient and non-sentient components.

Today we exist at a point in human evolution where rote tasks are increasingly abstracted from direct human involvement. Decisions about those tasks, however, are not completely abstracted from human oversight. Humans, therefore, may be viewed as part of the systems contrived to accomplish tasks in pursuit of improving their condition(s). We set forth the definition of a system as a set of things [elements] working together as parts of a mechanism [1]. We identify three fundamental elements of systems; 1) hardware or otherwise inanimate objects in physical space; 2) software or instructions guiding hardware operations; 3) human operators who implement instructions to accomplish tasks with otherwise inanimate objects.

Applying specific focus to effective system operations, we note that system knowledge is paramount to effective task execution. Traditionally, this knowledge implies intimate understanding of hardware and software system elements. Relatively recent advancements in biological sciences provide insight into the need for human operators to possess sufficient knowledge about themselves to perform increasingly complex functions as part of a system. In this paper we demonstrate the need for humans responsible for conducting space domain operational functions to attain a higher level of knowledge about themselves than the current average operator possesses. We recommend extending an already existing framework applied from the discipline of Behavioral Psychology, known as Crew Resource Management (CRM), from other domains to facilitate space domain specific knowledge acquisition and training.

Historical Imperative: Inception and Generations of CRM

Crew Resource Management (CRM) originated in the late 1970s after a series of catastrophic aviation accidents attributed not to mechanical failure, but to breakdowns in crew communication, coordination, leadership, and decision-making within the cockpit [2]. The 1977 Tenerife airport disaster, a collision of two Boeing 747s and the deadliest aviation accident in history, was a catalyst for the realization that a disciplined approach to dealing with human error was necessary. This recognition spurred a shift in thinking—from individual pilot error to systemic human factor shortcomings in team performance.

In 1979, the National Aeronautics and Space Administration (NASA) convened a workshop on resource management on the flight deck. It was here that the term “Cockpit Resource Management” was coined, emphasizing the need to address interpersonal communication, decision-making, and leadership among flight crews. CRM was conceived as a human factors training program designed to optimize team function and mitigate human error. It represented a profound shift in aviation training—moving beyond technical skills to focus on cognitive and interpersonal dynamics [2].

First Generation

The first generation of CRM training emerged in the early 1980s and focused primarily on psychological concepts such as interpersonal behavior, leadership, and communication. These programs were classroom-based, often led by psychologists, and aimed to increase awareness of human limitations and enhance teamwork. Although foundational, this generation of CRM lacked operational relevance and was seen by some flight crews as overly theoretical and lacking direct applicability to operational contexts [2]. These early programs also relied heavily on

role-playing and abstract scenarios, sometimes leading to skepticism among pilots and limited buy-in from operational leadership. Nonetheless, first-generation CRM succeeded in laying the groundwork for recognizing the critical role of crew coordination in flight safety [2].

Second Generation

During the mid-1980s, second-generation CRM shifted to a more operational approach, integrating CRM concepts directly into Standard Operating Procedures (SOPs). Focus shifted from psychological theory to applied practice. Programs emphasized scenario-based training and debriefs, incorporating real-world case studies of aviation mishaps, which helped crews understand and practice skills in simulated environments [2].

This generation saw CRM training become a regulatory requirement, with bodies like the United States Federal Aviation Administration (FAA) and International Civil Aviation Organization (ICAO) mandating it for commercial airlines. This institutionalization helped legitimize CRM in the eyes of operators and marked a transition toward organizational commitment to safety culture. Line-Oriented Flight Training (LOFT) became a central training method, simulating realistic flight scenarios that allowed crews to practice CRM principles in context.

Third Generation

Third-generation CRM expanded the scope of training beyond pilots to include other stakeholders such as flight attendants, maintenance personnel, and air traffic controllers. It adopted a systems-based approach that emphasized organizational safety culture, cross-functional coordination, and latent threat management.

At this stage, CRM was seen not just as a crew-level tool but as an enterprise-wide safety strategy. Programs like the University of Texas Threat and Error Management Model (UT-TEMM) were developed to support holistic error management across organizations. CRM also began addressing issues such as fatigue, organizational culture, and automation complacency—challenges increasingly relevant in complex modern aviation systems [2].

Fourth Generation

The fourth generation of CRM training, which began emerging in the 2000s, emphasized continuous learning and data-driven approaches. Training was now regularly refreshed and updated based on operational feedback, accident investigations, and human factors research. Crew behaviors were analyzed using observational tools like the Line Operations Safety Audit (LOSA), allowing for proactive threat identification and training refinement.

This generation also saw CRM's influence expand into non-aviation domains such as medicine, nuclear power, and firefighting. These industries recognized the parallels in human performance risks and adopted CRM-style training to reduce error and improve team performance under pressure. In addition to continuous learning, safety culture, and cross-disciplinary coordination also became central themes [2].

Fifth Generation

Fifth-generation CRM focuses on resilience engineering and adaptability in complex, high-uncertainty environments. It recognizes the limitations of procedural compliance and instead fosters adaptive capacity—the ability of teams to respond effectively to unforeseen challenges. Training emphasizes dynamic situational awareness, decentralized decision-making, and cultural agility.

As space operations become more automated and interagency in nature, this generation of CRM is especially relevant. It incorporates cross-disciplinary research from neuroergonomics, cognitive science, and behavioral economics, and integrates digital tools such as real-time analytics and AI-driven simulations. This generation is especially relevant to space operations, where dynamic and uncertain conditions require a high degree of operational agility [3]. This emphasizes the importance of shared situational awareness and redundancy in high-consequence environments like SDA.

Threat and Error Management (Fig 1 [2])

Originating from the third and fourth generations of CRM, Threat and Error Management (TEM) is a conceptual framework developed to proactively identify, address, and mitigate threats and errors that occur in complex operational environments. Originating from aviation safety research, TEM acknowledges that threats (external events or errors) are an inherent part of flight operations—and by extension, any high-risk domain. The framework encourages a shift from error avoidance to error management, focusing on resilience, recovery, and adaptability in real time [2][3].

Threats

A core tenet of TEM is that threats are inevitable and often outside the control of the operator (e.g., weather conditions, technical malfunctions, or time pressure). These threats can lead to human errors if not effectively detected or managed. The emphasis is on identifying these threats early and implementing defensive strategies such as briefings, checklists, and cross-checks to mitigate their impact before they escalate [2].

Errors

Errors, in contrast, are typically internal actions or inactions—such as miscommunication or incorrect data entry—that can compromise safety or performance. TEM classifies errors into categories (e.g., procedural, decision, proficiency, and communication) and tracks their outcomes (e.g., inconsequential, mitigated, or consequential). This structured approach helps organizations understand the nature of human error and develop interventions to reduce its recurrence [3].

TEM Applied

TEM is taught through scenario-based training, simulation, and observational tools like Line Operations Safety Audit (LOSA). Instructors use real-world examples to highlight common threats and errors, reinforcing best practices for recognizing and responding to them. Participants are trained to verbalize threats, confirm shared understanding, and use decision-making tools under time pressure. The model promotes a proactive rather than reactive safety culture [3]. Importantly, TEM supports a systems-thinking approach, encouraging organizations to identify latent conditions that contribute to threats and errors—such as poor scheduling policies, inadequate training, or hierarchical communication norms. This broader view extends accountability beyond the individual to the organizational level, reinforcing the role of leadership and culture in safety performance [2]. While TEM is most widely known for its application in aviation, it has also been adapted to healthcare, rail transport, and space operations. In these domains, it provides a valuable lens for understanding how teams cope with dynamic, high-stakes situations and how to build systems that support effective risk management and human performance [3]. Here are three real-world examples illustrating how Threat and Error Management (TEM) applies in practice:

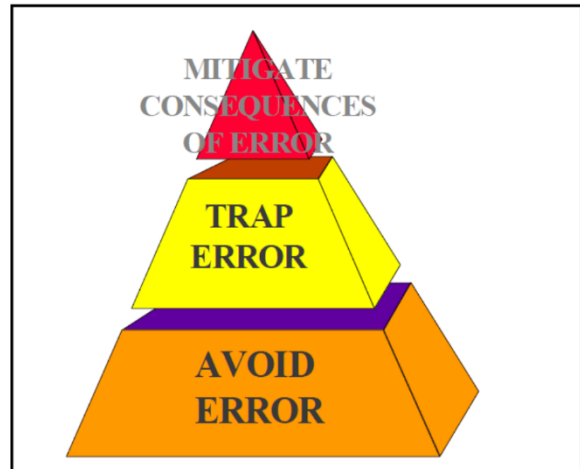


Fig 1

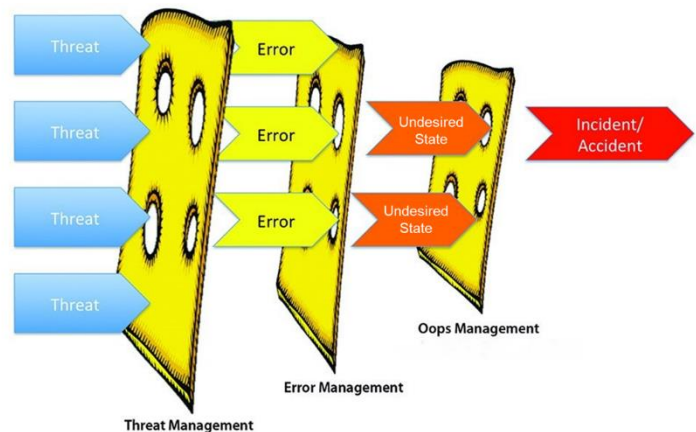


Fig 2: <https://sl.bing.net/jn24rIiyn6>

1. Aviation – Altitude Deviation During Approach

Threat: A sudden change in weather requires a last-minute runway reassignment.

Error: The pilot inputs the wrong altitude into the flight management system due to task saturation and miscommunication with the co-pilot.

Management: The co-pilot notices the discrepancy during a cross-check, initiates a correction, and the aircraft descends safely.

Outcome: Error mitigated, no operational consequence.

Explanation: This example highlights how CRM principles and proactive monitoring help manage threats and recover from errors before they escalate.

2. Healthcare – Surgical Sponge Count Discrepancy

Threat: A complex emergency surgery with multiple instrument trays increases procedural workload.

Error: The circulating nurse miscounts the sponges due to a distraction during the count.

Management: The team conducts a mandated post-op count and identifies the discrepancy. Imaging confirms a sponge was left inside the patient, which is removed in the same session.

Outcome: Consequential error but mitigated before serious harm.

Explanation: This illustrates how structured procedures and vigilance can catch human error even under pressure.

3. Space Operations – Satellite Collision Avoidance

Threat: A conjunction alert (potential satellite collision) is issued late due to ground system delay.

Error: The senior conjunction analysis operator initially misinterprets the alert severity due to fatigue.

Management: A junior team member verifies the trajectory and flags the risk, prompting a timely maneuver.

Outcome: Collision avoided; error caught through team redundancy.

Explanation: This example illustrates the application of CRM informed high performance team dynamics and behavior resulting in mitigating a potential mishap.

CRM in Other Domains

The aforementioned success of early CRM generations in aviation has led to widespread adoption of later generations in other high-risk sectors. One of the earliest adopters outside aviation was the healthcare industry. Operating rooms, much like cockpits, are high-stakes environments requiring rapid, coordinated decision-making under pressure. CRM has been integrated into surgical team training to improve pre-operative briefings, intraoperative communication, and post-procedure debriefs. This approach has led to measurable reductions in adverse events and improved team dynamics [4].

Routine inpatient and outpatient medical care has also benefited from CRM principles. In these environments, handoffs, care transitions, and interdisciplinary team communication are common failure points. CRM-inspired tools

such as SBAR (Situation-Background-Assessment-Recommendation) have been implemented to improve clarity and consistency. These adaptations help foster a culture of accountability and empower all staff to contribute to patient safety [5].

CRM principles have similarly been adopted in the nuclear power industry and in firefighting. In nuclear power plant control rooms, CRM emphasizes error detection, team leadership, and standard communication protocols during normal and emergency operations. In firefighting, CRM supports real-time coordination and situational awareness during rapidly evolving incidents. These implementations reflect CRM's core strength: enhancing human performance in complex, safety-critical systems [6].

2. Foundations in Psychology and Relevance to Space Operations

As previously stated here and in [2] CRM emerged from aviation's response to high-profile failures attributed to poor communication, leadership, and coordination. As generations of CRM matured, domains with dependencies on high performing teams began adopting and adapting CRM to their operational paradigms. Robert Helmreich's foundational work emphasized that effective CRM addresses both the technical and interpersonal aspects of operations [2]. In the previous section we emphasized that CRM includes structured training on leadership, decision-making, workload management, and team communication.

According to [7], human factors considerations must be systematically integrated into system operations and acquisitions. The FAA emphasizes a systems approach, recognizing the interplay between human capabilities and technological systems. CRM, as applied in aviation, embodies this approach by enhancing team coordination, communication, and decision-making. These elements are equally critical in space operations, where miscommunication or poor coordination can result in catastrophic outcomes. Domains that share common dependencies on high-performance teams benefit from applications of CRM principles. High-performance teams are defined here as a group of people regarded as tight knit, focused on their goal and having supportive processes that will enable any team member to surmount any barriers in achieving the team's goals. We offer a second, complementary description of high-performance teams identified as a group of people with specific roles, complementary talents and skills, aligned with and committed to a common purpose, consistently showing high levels of collaboration and innovation to produce superior results. In high-performance team settings, CRM tenets extinguish radical or extreme opinions that could be damaging.

Space operations present similar complexities consistent throughout high-performance team dynamics: high workload variability, extended periods of passive monitoring, and sudden transitions to emergency response. These contexts mirror the conditions that Helmreich and others found to be most error-prone in aviation and medicine as presented in [3]. CRM principles are essential to mitigate the human errors associated with information gaps, high power-distance cultures (i.e. organizations with stratified seniority and/or rank structures requiring active controls to enable clear communications throughout high performance team members), and unspoken team assumptions.

Scientific Foundations

CRM concepts are rooted in the study of human behavior dating to the early twentieth century. Key foundational concepts are presented via a review of pertinent literature, published research, and certain key terms.

Human Performance as a Function of Stress

The Yerkes-Dodson law is a psychological principle that maps the empirical relationship between arousal (commonly associated with stress) and human performance. Originally developed in 1908 by psychologists Robert Yerkes and John Dodson, the theory asserts that performance improves with increasing arousal, but only to a point. Once arousal exceeds an optimal level, further increases lead to a decline in performance. This dynamic is often visualized as an inverted U-shaped curve—the Yerkes-Dodson curve—which has become a foundational concept in human factors and applied psychology (Fig. 3) [8].

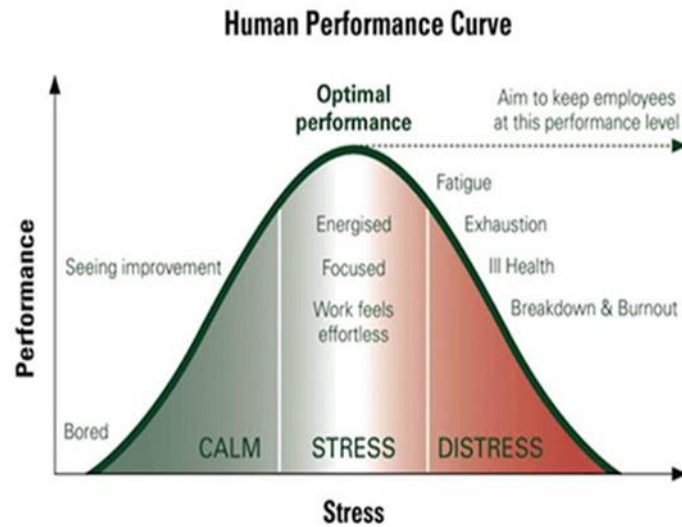


Fig. 3 Yerkes-Dodson Curve

Different tasks demand different levels of arousal for optimal performance. For simple, repetitive tasks, higher arousal or stress levels may enhance performance. However, for complex tasks requiring memory, decision-making, or fine motor skills—such as those often encountered in aviation, healthcare, or space operations, excessive stress can overwhelm cognitive resources. In these situations, excessive stress can result in information overload, impaired decision-making, decreased situational awareness, and prolonged response times to key stimuli.

The Yerkes-Dodson Law applies to all high-reliability domains where high-risk, high-consequence environments are common. In these environments high-performance teams comprised of human individuals are expected to execute tasks in the face of challenging and often unforeseen circumstances. Certain case studies in aviation and emergency response illustrate both the benefits and pitfalls of stress loads on high performance teams. For example, moderate stress levels can improve reaction time and vigilance during an inflight systems failure, but excessive stress may lead to tunnel vision or premature decision-making, which compromises safety. This empirical framing reinforces the need to manage operator stress through system design, team dynamics, and training protocols.

Subsequent research and analysis into the Yerkes-Dodson Law integrates with broader human performance literature, including [9] and [10]. In her analysis of the Yerkes-Dodson law, [8] highlights how vigilance loss over time—an issue also addressed in [9]—is exacerbated by arousal levels that are either too low (boredom and underload) or too high (anxiety and overload). The Law’s implications closely align with the findings in [9], [10] and [11] discussed in subsequent sections of this paper. These connections underscore the importance of designing operations and training environments that sustain operators within the “optimal arousal zone” proposed by Yerkes and Dodson.

The same arousal-performance dynamics contributing to vigilance decrement as noted in [9] reinforce the inverted-U hypothesis of Yerkes-Dodson. Specifically, insufficient or excessive arousal can degrade attention and performance over time, regardless of operator skill or motivation. This shared insight points to the value of applying arousal management principles in training design and real-time performance support tools.

Ultimately, the Yerkes-Dodson model offers both diagnostic and prescriptive value for human systems integration. Reference [8] recommends incorporating stress monitoring, adaptive user interfaces, and periodic recovery intervals to keep operators within the “optimal arousal zone.” In this way, the curve complements and informs CRM and TEM strategies, providing a tool for enhancing operator resilience, safety, and mission success.

Intersections of Automation, Vigilance, Skill Retention, Error, and Failure

One of the key sources of poor vigilance is synonymous with complacency. Complacency is a major indicator of low situational awareness. Low situational awareness correlates to states of increased operational risk, especially during protracted periods of low stress, arousal, and/or cognitive loading. Over-reliance on automation and abstraction of cognitively demanding tasks results in diminished capacity for individuals and human teams to react to off-nominal conditions when they occur. The dynamic nature of operations dictates the unpredictability of when and how off-nominal conditions, and a corresponding sudden onset of cognitive loading, manifest themselves. Once a threshold of risk factors aggregation is reached, undesirable irreversible end states also known as mishaps tend to occur.

Automation – Definitions and Implications

[12] introduces a comprehensive philosophy of human-centered automation applied to operating aircraft. Observations and contentions are directly germane to core principles which inform the applicability of CRM to space domain operations. As in [12] we affirm, that automation should support, not supplant, human operators. We leverage a conceptual framework in which automation serves as a cooperative partner in operational tasks, emphasizing shared platform/system control, transparency, and predictability. The goal for system design is to build systems that enhance operator performance and judgment, while minimizing degradation situational awareness or cognitive engagement. Reference [12] warns, and history has consistently shown, that poorly designed automation can create hazards by displacing human control without sufficient feedback or override capability. Additional reviews of human operator focused research, described in subsequent section of this paper, affirm these assertions.

We note that Space Domain operations are subject to points made in [12] that, historically, automation has significantly improved reliability and workload distribution but at the cost of increasing cockpit/console complexity and potentially reducing operator engagement. By leveraging the three functional categories of aviation automation found in [12]: control (e.g., automated platform control), information (e.g., system status displays), and management (e.g., mission planning systems), we begin to identify key nodes that are necessary for sustaining operator arousal. System elements supporting each functional category must keep operators informed, involved, and in command. These are foundational principles of a “human-centered” approach to system design. We highlight these principles here to emphasize the importance of training scoped to the systems and conditions operators routinely interact with.

Next, we draw attention to attributes of effective automation found in [12], which include predictability, comprehensibility, accountability, and error tolerance. We draw a parallel path for cognitive engagement of space domain operators from the advocacy found in [12] for designs that enable pilots to understand the logic of automated actions and recover from failures when they occur. Importantly, [12] also emphasizes that automation must be able to monitor human performance, just as humans monitor automation—a reciprocal relationship that mirrors later concepts in crew resource management (CRM) and threat and error management (TEM).

The framework in [12] intersects with interpretations of the Yerkes-Dodson law made in [8], mapping performance to arousal levels. Automation that removes humans from meaningful control can lower arousal below the performance threshold, while ambiguous or poorly integrated systems can overwhelm them. These parallels further affirm the need for human-centered automation as a design philosophy grounded in behavioral science.

We turn to [13] which outlines the evolving definitions, classifications, and implications of automation in flight systems, with relevance across other industrial settings. Reference [13] traces early definitions, noting the lack of consistency until the widely adopted version found in [12]: “Automation is a process that controls a function or task without human intervention.” This foundational concept sets the stage for analyzing how automation, viewed as a continuum, influences operator roles, problem solving skills, cognitive workload, and system safety. This definition is immediately extensible to other domains where automation is applied to augment human operational performance.

Key frameworks presented in [14] and [15] further define automation in levels or layers of human-system interaction. These range from manual control to full autonomy, with intermediate stages like decision support and monitored AI. Reference [13] points out that different aviation professionals perceive automation in just three operational levels—manual, partially automated, and fully managed by flight management systems—underscoring the divergence between theory and practice. Universal themes are identified, specifically concerning human factors such as loss of manual skill, complacency, and degraded performance in high automation scenarios.

As shown in [13], empirical research supports these concerns. Pilots in simulated environments showed reduced cross-checking behavior and increased task-unrelated thought when flights were uneventful or heavily automated. Findings also suggest that while automation supports instrument monitoring and manual control, it may impair cognitive skills like navigation decision-making. This degradation is often unnoticed until an anomalous state forces a quick transition to manual control, compounding risk due to unfamiliarity or atrophy of critical skills.

Given observation of operators under benign controlled conditions, we turn to real world examples for validation of initial observations in [13]. A subsequent accident analyses compilation in [16] identifies recurring failure patterns associated with automation misuse and misunderstanding. Conceptual clarity is added to what “automation” actually means, and how different forms, as defined in [12], carry distinct risks. Likewise, a review of vigilance in [9] provides explanations for why prolonged monitoring under automation induces cognitive fatigue, reducing error detection. Together, we leverage a cohesive argument, applying it to the space domain: automation is not a binary or neutral design choice—it reshapes operator roles and safety dynamics at every level. We therefore assert the importance of training operator self-awareness as a key element to system operations.

[13] re-enforces themes found in [8] on the Yerkes-Dodson curve, showing how inappropriate levels of automation can push operators into cognitive underload or overload. Just as the Yerkes-Dodson model suggests a peak performance zone tied to optimal arousal, [13] implies that system design must calibrate automation levels to sustain operator engagement and resilience. Conclusions drawn from multiple sources advocate for adaptive interfaces and contextualized training as countermeasures to the risks of disengagement and performance collapse.

We echo warnings from [13] against treating automation as a panacea and advocate for user-centered design that considers mental models, feedback mechanisms, and training needs. Automation, while intended to assist, introduces its own category of risks—ones that escalate when operators are removed from the loop or left without sufficient insight into system behavior. Through historical context, taxonomy, and cross-sector analysis, we leverage cautionary insight and design-oriented perspectives for integrating automation into complex domain operations.

Accident Analysis and Human Error

Human error is ubiquitous and unavoidable. Few domains provide more acute illustrations in recent history of the consequences of human error than aviation. Exploring the root causes of aviation mishaps, [16] focuses specifically on systemic and design-related contributors to human error. Reference [16] points out that automation, while reducing routine workload, can obscure an operators’ understanding of system status and reduce their ability to intervene effectively when anomalous situations occur. Citing several high-profile aviation accident case studies spanning a 41-year epoch, [16] demonstrates how over-reliance on automation, poorly designed human-machine interfaces, and procedural error led to operator loss of situational awareness and, ultimately, undesired irreversible end-states, also-known-as accidents or mishaps.

In an expanded study of specific cases, [16] highlights aviation accidents where automation played a significant role, aiming to uncover patterns and vulnerabilities in flight deck systems, analogous to space domain operator consoles and control systems. Drawn from Federal Aviation Administration (FAA) and NASA data, there are a subset of incidents compiled in which pilots failed to properly manage automated systems, leading to catastrophic outcomes.

These cases include the Asiana Flight 214 crash in San Francisco and multiple Airbus A320 incidents. We note a recurring theme that system design often obscures status cues or provides inadequate feedback, making error detection difficult. This theme echoes concerns in [12] about information overload and the illusion of system reliability. Reference [16] argues that the interface design, documentation, and training frequently fail to equip pilots to handle these nuances. Notably, recurring issues included mode confusion, incorrect settings, complacency, and over-reliance on automation during critical operational phases. Reference [16]

A major contribution we draw from case study analysis in [16] is a detailed taxonomy of automation issues compiled from various studies, including [17] and [18]. These problems are categorized into design flaws, usage challenges, and organizational shortcomings—highlighting the systemic nature of automation-induced failures. We incorporate this taxonomy into the list of tools available for space domain operator training development and sustainment. Reference [18]

[16] also emphasizes the importance of contextualizing human error as the result of upstream organizational and technological decisions. When analyzing mishap root causes, Dr. Gawron advocates for a broad assessment that includes latent system deficiencies and procedural rigidity. A key theme extensible to the space domain is the need for improved data collection during normal operations to identify and address vulnerabilities before they manifest as undesired irreversible end-states. Gawron points out that many errors are repeat occurrences of previously known issues—warnings that were inadequately addressed or dismissed as anomalies. Later we illustrate how repeat occurrences of previously known issues can compound, increase operational risk, and eventually lead to mishaps which follow a subtle change in domain conditions [19].

One solution proposed in [16] is embedding structured observation and reporting protocols to close the learning loop. In practical terms this takes the form of a feedback mechanism, necessary for training to keep pace with the dynamism which inevitably occurs in any domain. Consequently, Gawron advocates for cultures of continuous learning within high-reliability organizations. Conclusions in [16] stress that mishap prevention must be proactive, supported by robust human factors engineering and organizational awareness, rather than reactive, waiting for failure to drive reform.

Failure Detection Research

Given the desire to prevent mishaps, we turn to investigations of human operators' ability to detect system failures in both manual and automated environments. Reference [11] presents findings from experimental studies, simulations, and field observations that show how detection times increase, and performance degrades as automation increases. Reference [11] also identifies key challenges: reduced engagement, passive monitoring, and diminished proprioceptive cues, all of which erode situational awareness and contribute to missed or delayed recognition of system anomalies.

One of the central findings of [11] is the “out-of-the-loop” performance problem, where operators disengaged from active control struggle to re-engage effectively when a system deviates from expected behavior. Reference [11] points out that even highly trained professionals are vulnerable to cognitive lags during transitions from passive to active roles, particularly under time pressure or ambiguous cues. These delays increase the likelihood of inappropriate corrective actions, i.e. human error.

To counteract these effects, [11] proposes enhanced training for pattern recognition and cue utilization, including the use of realistic simulations that introduce ambiguous or degraded information. Reference [11] also highlights the value of multi-modal displays and collaborative monitoring (i.e., multiple team members verifying system health), both of which improve early error detection rates. The presence of these factors in existing CRM training is no coincidence.

This research is presented to address one of the specific failure modes cited in accident analysis: operators' inability to detect and react to system anomalies in time. We therefore advocate for those responsible for operator training in a high-reliability domains dissect the cognitive mechanisms behind failures, recommend methods for mitigation, and integrate them into formal training via active feedback mechanisms.

High-reliability operations are therefore likely to benefit from the development of adaptive automation and vigilance support systems that adjust the level of operator engagement based on task criticality. Emphasis on early detection and seamless transition between control modes reinforces his broader advocacy for system designs that prioritize human cognitive strengths rather than replacing them. We now draw a conclusion that adaptive automation and vigilance support is necessary to provide sufficient cognitive load for effective high-reliability operations.

Skill Retention Research

Next, we pose the question, how and to what levels must operators be stimulated to attain and sustain necessary levels of vigilance and proficiency? In her compilation of skill retention research, [10] examines how cognitive and psychomotor skills degrade over time, especially in low-tempo or highly procedural domains. Using synthesized data from military, medical, and aviation studies, [10] shows that while basic procedural skills can remain stable with minimal practice, complex or emergency-related tasks degrade rapidly without repetition.

In [9] skill retention is categorized into “retention intervals” and task types, explaining that emergency and decision-making tasks (e.g., abort procedures or engine failures) decay much faster than routine operations like checklist usage. Her analysis emphasizes that both the frequency and fidelity of practice are critical in maintaining proficiency, particularly for tasks not used daily but crucial during high-stress situations.

Analysis in [10] emphasizes the importance of training interventions such as overlearning, distributed practice, and booster sessions, which help mitigate skill decay. Consequently, [10] supports the use of low-cost simulators and scenario-based refreshers as practical methods to reinforce skills over time. These interventions and training methods are facets of existing CRM courses in other domains. She also notes, and we highlight, that skill decay is not uniform across individuals—experience, cognitive load, and confidence all influence how quickly skills are lost.

Understanding skill retention, and the non-uniformity of human retained skills are key to sustaining effective operations because employable skills are often prerequisites for recognizing abnormal indications and managing failures. Understanding skill retention informs domain training enterprises, building on the systemic perspective from accident analysis, highlighting how gaps in refresher training contribute to accidents by eroding operators' readiness.

As stated in [10], operational organizations must treat skill retention as a continuous process, not a one-time training milestone. We therefore advocate for embedded assessment loops and performance tracking to personalize retraining and optimize long-term operational safety performance. In application, we leverage knowledge of human skill retention to guide training design, modality(ies), and cadence.

Vigilance Research

Diving deeply into the psychology of vigilance, the concept of sustained attention during monotonous tasks, we present a review of four dominant theories of vigilance decrement from [9]: Signal Detection Theory, Arousal Theory, Expectancy Theory, and the Vicious Circle Hypothesis. Originally rooted in WWII-era radar monitoring and directly applicable to routine space operations functions, the research illustrates how operator performance deteriorates over time, especially when off-nominal signals are rare, ambiguous, or low in intensity.

I. Signal Detection Theory (SDT)

Signal Detection Theory (SDT) is a foundational model for understanding vigilance tasks, particularly in environments where individuals must monitor for rare or subtle signals (e.g., aircraft radar, air traffic control, or

satellite tracking). According to SDT, an observer must continuously distinguish between signal (important, task-relevant stimuli) and noise (irrelevant background stimuli). The theory introduces concepts such as “hits,” “misses,” “false alarms,” and “correct rejections,” and considers an observer’s *sensitivity* (ability to discriminate signals from noise) and *response bias* (tendency to be conservative or liberal in reporting detections). The theory assumes that decision-making is influenced by the perceived probability of signal presence and the consequences of incorrect responses.

In [9] we see emphasis that SDT explains a key challenge in vigilance tasks: even when individuals are well-trained, their ability to detect rare signals deteriorates over time due to cognitive fatigue or low signal frequency. This makes SDT highly applicable to space domain awareness and satellite monitoring, where events like debris conjunctions are rare but critical. SDT also interfaces with automation issues; over-reliance on automated systems can degrade the operator’s sensitivity and make them less responsive when manual intervention is suddenly required. This theory underpins discussions in [11] on real-time failure detection and explains why missed anomalies are so common in automated, passive-monitoring environments.

II. *Arousal Theory*

Arousal Theory proposes that vigilance performance is directly tied to the operator’s level of physiological or mental activation. According to this theory, low arousal—often caused by monotonous, unstimulating tasks—leads to decreased attention, slower reaction times, and higher error rates. In contrast, moderate levels of arousal can enhance performance by increasing alertness and responsiveness. Gawron cites experimental studies in which arousal was manipulated using external stimuli (e.g., threat of shock, noise, or caffeine), with results showing short-term boosts in detection performance.

However, [9] notes that artificial arousal-enhancement techniques have limited utility in operational settings due to practicality and sustainability. More importantly, excessive arousal—particularly under stress—can lead to anxiety, tunnel vision, or premature decisions, as outlined in [8]. Thus, Arousal Theory interacts closely with both SDT and [8], showing that vigilance must be supported by balanced system design and environmental stimuli. Arousal Theory reinforces Gawron’s broader theme: that optimal vigilance occurs in a cognitive “sweet spot,” and system design should help maintain operators within it.

III. *Expectancy Theory*

Expectancy Theory explains vigilance performance in terms of the operator’s expectations about when, where, or how often a signal will appear. It posits that individuals unconsciously (or consciously) form expectations based on the frequency and timing of past signals. When signals are infrequent or unpredictable—as they often are in space monitoring or control room operations—expectancy is low, and attention tends to decline. Research cited in [9] shows that when participants are told to expect a signal at certain intervals, their performance improves in those time windows, even if the signal itself doesn’t change.

This theory has clear implications for operational environments with rare but critical events. If operators rarely encounter real alerts, their expectation for action decreases, which in turn reduces attention and vigilance. Findings in [9] suggest that training and simulation can help offset this by periodically injecting “false positives” or simulation drills to elevate expectancy levels. This theory complements SDT by providing a behavioral and cognitive explanation for decreased sensitivity in long-term monitoring tasks. It also connects to skill retention theory from [10] by explaining how expectation interacts with task rehearsal and memory.

IV. *The Vicious Circle Hypothesis*

The Vicious Circle Hypothesis proposes that errors and low performance in vigilance tasks can create a self-reinforcing feedback loop. When an operator begins missing signals due to fatigue, low arousal, or other factors, they may become demotivated or anxious, which further decreases performance. This decline can result in increased

stress or negative self-assessment, making it harder to recover attention or confidence. This theory helps to explain long-term degradation in vigilance, particularly in isolated or high-pressure roles where feedback is delayed or absent [9].

This hypothesis is particularly important in understanding skill decay and vigilance in space domain operations. Operators who go long periods without meaningful engagement may begin doubting their capabilities, which in turn reinforces cognitive disengagement. Reference [9] links this to automation complacency and the broader implications of poor system design. When coupled with limited practice and inadequate feedback, as discussed in [10], the vicious circle can erode both performance and morale. It reinforces the need for built-in feedback, performance monitoring, and recurring training to break the negative cycle.

Analysis of vigilance explains why operators fail to notice automation errors or detect anomalies—a problem elaborated in those documents. Additionally, [9] challenges the assumption that vigilance tasks are boring and mentally underloading, citing research that shows they are in fact cognitively draining and emotionally taxing. It supports advocacy for designing systems and training protocols that sustain attention and minimize operator fatigue. Conclusions in [9] show that understanding vigilance is crucial to improving human-machine teaming. We echo calls in [9] for further applied research that aligns detection theory with real-world monitoring environments, especially as automation reduces the frequency—but not the importance—of human intervention. We derive the following implications for CRM/TEM training program development, informed by these vigilance theories.

- I. Enhance signal salience in simulations; adjust signal-to-noise ratios; include practice with ambiguous cues. Teach operators to calibrate their response bias based on task risk and signal frequency.
- II. Design tasks with varied pacing and challenge levels to maintain optimal arousal. Use moderate stress-inducing scenarios (e.g., time pressure or uncertainty) in training to replicate real conditions. Link arousal management to the Yerkes-Dodson curve to balance task intensity and performance.
- III. Include random and expected signal events during training to raise attention thresholds. Reinforce periodic expectation with feedback and realistic alert injection. Use drills to establish accurate mental models of signal frequency.
- IV. Include immediate feedback mechanisms in both simulations and live environments. Reinforce early success to build confidence. Design recovery protocols to restore engagement after performance dips.

Automation Guidelines

The final document included in our review shifts from accident analysis to offering proactive design and training recommendations aimed at mitigating the risks posed by automation. Pivoting, [20] synthesizes decades of aviation research to produce guidelines that consider technical, cognitive, and organizational factors. Reference [20] clearly emphasizes the need for systems to be intuitive, provide reliable feedback, and be designed around the human operator's strengths and limitations.

One of the report's key features is a Human Factors Checklist, which draws from [21] and considers cognition, social dynamics, ergonomics, and environmental stressors. The checklist is intended to guide both system designers and accident investigators. Also referenced in [20] are cognitive mismatch issues between operators' mental models and a system's actual behavior, a disconnect she argues which contributes to many automation-related incidents.

Survey data from pilots show that while they generally appreciate automation, they also report confusion about mode behavior, limited feedback, and reduced engagement. A warning emerges of the risk of over-reliance on automation, which leads to calls for training that fosters deeper understanding of system operational functions, not just procedural knowledge. Ultimately, [20] advocates for a human-centered approach to automation—one that prioritizes clarity, control, and continuous training. We echo warnings in [20] that without learning from past failures, future systems may amplify rather than mitigate risks.

3. Historical Context

A historical example that underscores the importance of integrating CRM tenets is the NASA Apollo 1 disaster in 1967 [22]. This tragic event, which claimed the lives of three astronauts, marked a pivotal moment in the history of human space exploration. It revealed significant flaws in organizational culture, system design, and communication. Among the contributing factors were faulty spacecraft design, inadequate safety protocols, and a lack of attention to the complexities of human interaction with technology [23]. These failures highlighted the urgent need for a systems approach that better integrates human factors into design and operations. This lesson remains relevant as AI becomes increasingly embedded in critical systems. A key lesson to emerge from the Apollo disaster was the importance of a safety-first organizational culture. As an inherently human influenced paradigm, organizational culture is a function of human choices, standards of behavior, training, individual and group heuristics. NASA's preoccupation with schedule pressures and cost-cutting led to systemic oversights, including failure to account for the human factors that could affect the mission's success [24].

Similarly, the 2009 collision between the Iridium 33 and Cosmos 2251 satellites exposed significant flaws in inter-agency communication and the limitations of automated systems in space operations [19].

Case Studies: Lessons from Past Space Domain Failures

Historical failures in space operations offer critical insights into the complexities of human interaction and organizational dynamics, especially as dynamism in space operations increases. Historical incidents such as Apollo 1, Apollo 13, STS-51-L, STS-107, and the Iridium 33-Cosmos 2251 collision underscore the relevance of CRM principles. In Apollo 13, team cohesion and communication under stress were critical to survival [25]. Conversely, the Challenger and Columbia disasters revealed failures in information sharing and leadership that CRM training could have addressed [26], [27], [28]. These case studies reinforce the importance of behavioral competencies in technical operations. In so doing, each event highlights unique challenges and provides lessons that can inform future approaches to human-AI collaboration, emphasizing the importance of communication, organizational culture, sustaining vigilance, and Human-Machine system redundancy.

NASA Apollo 1 Disaster: Communication and Cultural Challenges

The Apollo 1 fire in 1967, which claimed the lives of three astronauts during a pre-launch test, serves as a stark reminder of the consequences of inadequate communication and systemic organizational flaws [22]. A key factor contributing to the tragedy was the lack of adequate feedback channels between engineers and astronauts. Critical design flaws, such as the highly flammable cabin materials and a cumbersome inward-opening hatch, were not sufficiently addressed due to poor communication across the project's hierarchical structure [23]. This failure stresses the need for transparent and iterative feedback mechanisms in high-stakes environments where human lives are at risk.

Organizational culture also played a significant role in the Apollo disaster. NASA's intense focus on meeting aggressive schedules and maintaining the United States competitive edge in the space race often overshadowed safety considerations. This "schedule over safety" culture created an environment where dissenting voices were minimized, and critical safety concerns were dismissed [25]. Addressing similar cultural pitfalls is crucial in the current era of AI integration, where transparency and collaboration between human operators and AI systems are paramount. As the space industry moves forward, it is crucial to learn from these historical failures and implement systems that integrate human factors and AI to enhance safety, decision-making, and overall mission success [15].

This failure emphasized the need for improved communication between engineers and astronauts, as well as better integration of human factors engineering into system design, ensuring that both human and technological capabilities were aligned to prevent catastrophic outcomes [2].

NASA Apollo 13: ‘A Successful Failure’

The Apollo 13 mission in 1970 is one of the greatest space domain exemplars of the importance of staying calm and methodical in a crisis. When an oxygen tank exploded, the crew faced a life-threatening scenario in cislunar space [25]. Their ability to remain composed, communicate clearly, and collaborate with the mission control team on the ground transformed a potential disaster into what’s now called a 'successful failure.'

Effective communication and decisive leadership under pressure, all principles at the heart of Crew Resource Management contributed to the safe return of the crew. In contrast with the errors that led to the Apollo 1 mishap, during Apollo 13 NASA engineers and astronauts worked together seamlessly, prioritizing tasks, sharing critical information, and executing solutions step by step.

A visible Cycle; NASA STS-107, Space Shuttle Columbia and STS-51-L Space Shuttle Challenger

Critical decisions made under pressure, when coupled with poor communication can induce a sequence of errors that lead to catastrophic outcomes. The loss of Challenger in 1986 and Columbia tragedy in 2003 painfully illustrate this point. Challenger was lost shortly after liftoff when a cold-soaked O-ring at the base of a solid-rocket booster cracked, exposing the rest of the spacecraft to high temperature flames, ultimately resulting in a catastrophic explosion [26], [27]. During the Columbia mission, insulation debris from the Space Shuttle External Fuel tank struck the leading edge of one of the orbiter’s wings during launch [28]. Unfortunately, in both cases warnings from engineers about potential risks and actual damage were not effectively escalated. These communication breakdowns contributed to the loss of the orbiters and their crews.

Both events illustrate two key points that our CRM foundation builds upon. First, CRM fosters an environment where every voice, regardless of rank, can raise concerns and be heard. Structured debriefs and clear communication protocols ensure that critical issues are addressed to mitigate risks.

Second, human knowledge, skill, expertise, and learned behaviors are perishable when not sustained. This is true for individual as well as organizational attributes. During periods of high emphasis where CRM concepts are embraced, we see a pattern of increased vigilance guarding against behavioral patterns considered risky that can increase the likelihood of erroneous human decisions. These cases clearly show that the opposite is also true.

If open communications principles had been fully realized during the STS-51-L and STS-107 missions, the outcomes may have been different. For future missions, especially as we venture deeper into space, applied CRM principles can ensure that decision-making is robust, inclusive, and focused on safety. We also note that cultural foundations that sustain implementation of these principles can prevent an oscillating cycle between high and low attention that is punctuated by mishaps occurring during low emphasis periods.

Iridium 33 - Cosmos 2251 Satellite Collision: Automation and Coordination Failures

The 2009 collision between the Iridium 33 and Cosmos 2251 satellites illustrates the challenges of automating decision-making in space operations. This incident, which destroyed both satellites and created thousands of pieces of space debris, highlighted the limitations of existing tracking algorithms. Automated systems failed to provide accurate collision predictions due to gaps in data reliability and a lack of sophisticated models capable of accounting for complex orbital dynamics [29]. This shortcoming demonstrates the need for adaptive AI (Artificial Intelligence) systems capable of real-time data integration and predictive analytics to enhance decision-making in space operations.

Inter-agency communication also emerged as a critical issue in the Iridium-Cosmos collision. The lack of coordination between the satellite operators, Iridium Communications, and the Russian Ministry of Defense exacerbated the situation. Data sharing and collaborative risk assessment protocols were insufficient, leading to missed opportunities for intervention. The lack of robust human oversight in satellite collision avoidance systems

underscored the need for integrated frameworks that combine human factors expertise with advanced AI to ensure that automated systems can be used effectively without over-relying on them [30], [31], [32].

This case underscores the necessity of fostering inter-organizational cooperation and developing global standards for space traffic management [33]. Incorporating these lessons into AI-enabled systems could ensure smoother communication and more effective collaboration between human and machine actors.

Additional Lessons Learned from Apollo 1 and Iridium 33

Both the Apollo 1 disaster and the Iridium-Cosmos collision point to the necessity of developing frameworks that integrate human factors and AI capabilities within systems engineering. These events reveal the challenges of human-AI interaction in critical systems, where human decision-making and technological processes must work harmoniously to prevent catastrophic failures. Effective system designs should prioritize communication, trust, and collaboration between human operators and automated systems, ensuring that both can function optimally within complex and high-pressure environments.

4. Pathway to Integration into Space Domain Operations

As directly stated in [2], “human error is ubiquitous and inevitable.” Humans are part of operational systems, not abstract observers or semi-active participants. As biological creatures, human behavior adheres to biologically defined cyclic activity patterns. While complex, human behavior is definable, identifiable, measurable, and to an extent, predictable. Executing an Observe Orient Decide Act (OODA) loop through a challenging situation that culminates in a desirable outcome depends on timeliness. Timeliness depends upon systems knowledge and proficiency that underpin sound decision making. Because humans are part of the operational systems in use, they must be trained effectively about themselves, in addition to the platforms, payloads, and material resources that they use to perform tasks.

While automation and AI tools reduce certain burdens, as shown in our literature review, they simultaneously increase the responsibility on human operators to detect and correct failures under high-stress, low-stimulation, or ambiguous conditions. The increasing technological complexity of space missions necessitates enhanced team coordination and advanced human performance strategies; traditional training methods are no longer sufficient to meet these demands.

Recent advancements in behavioral science provide modern tools and methodologies that enable better assessment and integration of behavioral psychology principles into training programs. The current gap in training—specifically the absence of comprehensive CRM training, and operational leadership development, in the space domain—presents a unique opportunity to make a significant and meaningful impact. CRM has demonstrated its value in aviation, healthcare, and other high-reliability sectors as a mechanism to enhance safety, coordination, and performance. To date, space domain operations, particularly unmanned spacecraft and SDA missions—have not adopted CRM in a standardized fashion.

Manned spaceflight operators, such as astronauts and mission support personnel (e.g., mission control console operators), do receive a form of CRM training. Historically, we identify episodic instances where certain NASA teams, like the Conjunction Assessment Risk Analysis (CARA) team, have received CRM-related training [34]. Their operational manuals contain valuable information applicable to space domain awareness operator training that can be leveraged for space domain operations CRM training.

Despite these isolated cases, CRM is absent from other space domain operations training. There is a lack of standardization, as no comprehensive CRM training has been specifically tailored for satellite and SDA operations. Operators in these functions do not receive consistent guidance on the best practices for leadership, teamwork, and communications. Existing training programs tend to focus on individual performance, emphasizing individual

operator skills while neglecting the importance of team dynamics and effective communication strategies that are crucial for mission success.

While there is substantial scientific data on the technologies and tools used by operators, there is a notable gap in applying behavioral psychology principles to improve team dynamics and human performance in space operations. This discontinuity fails to equip operators with tools to recognize and mitigate human errors stemming from psychophysiological factors such as those alliterated in [2] – [13], and myriad others.

5. Opportunities

Since human error is unavoidable, it is a significant contributor to mission risk in operations of complex systems, and nowhere is this more consequential than in the domain of space operations. While automation and artificial intelligence tools have reduced certain cognitive burdens, they have also heightened the demands on human oversight during anomalies and emergencies. Historically, aviation and manned space programs have adopted Crew Resource Management (CRM) to mitigate these risks, yet satellite and space situational awareness operations lack standardized CRM training. We outline the necessity and benefits of CRM in these emerging operational contexts.

Research

Research on cockpit communications reveals that language structure, workload, and stress directly affect communication effectiveness, team coordination, and performance [3]. Failures in crew interaction—particularly unclear or indirect language—were frequently associated with errors and higher perceived workload. This highlights the necessity for studying operations in space environments where communication clarity under pressure is critical.

Moreover, culture—national, organizational, and professional—plays a significant role in error management and team behavior. Current space domain operational organizations possess structures that reflect high levels of power-distance as a function of rank, seniority, perceived, and actual authority. High power-distance cultures, exemplified by highly stratified organizational structures where position, experience, and powerful personality dynamics often inhibit subordinates from speaking up, while professional cultures can promote overconfidence and resistance to acknowledging error [2]. CRM counters these effects by institutionalizing norms of assertiveness, feedback, and shared situational awareness.

CRM related research has stagnated; since the retirement of leading researcher Robert Helmreich in 2011, advancements in CRM research specific to this domain have significantly slowed. This stagnation exacerbates the gap in training methodologies that do not address the evolving complexities of modern space missions. Additionally, there is a noticeable deficiency in leadership development at all levels – junior, middle, and senior team members. Although there may eventually be standardization in team leader training, to address the risks we highlight, it must include a focus on individual performance and behavioral forces. Poorly scoped leadership development leaves overall team optimization and efficacy lacking at a time that the need for operational readiness is increasing.

Training

To set the stage for appropriate training development, we point to guidelines promulgated in the aviation domain. Reference [35] highlights how interruptions and distractions lead to deviations from standard operating procedures (SOPs), often resulting in accidents. CRM offers structured techniques to recognize, prioritize, and manage such disruptions effectively. Similarly, [36] provides a roadmap for developing tailored CRM procedures integrated into SOPs, emphasizing instructor and crew training, assessment, and continuous feedback.

The approach we present introduces space domain-specific CRM training that includes leadership development modules. This approach engages in rigorous scientific research prior to the development and implementation of training. This involves conducting comprehensive studies and human subject research with domain-specific operations personnel across military, intelligence, civil, and commercial sectors to ensure that the training is

grounded in real-world experiences and needs. This includes activities ranging from ground-based sensor management, centralized scheduling and tasking, coordination efforts, to spacecraft control and operations involving satellites and space situational awareness (SSA) sensors/sensor networks. Additionally, we recommend multimodal data collection methods that include observations of actual operations, participation in exercises, and the use of controlled training events specifically designed to evaluate and stress-test space domain operators. This comprehensive approach allows us to identify the most critical CRM principles applicable to space operations.

An effective training syllabus is subdivided into two main areas to address the specific needs of different operators. The first area focuses on Spacecraft Operations, concentrating on CRM principles relevant to satellite control and mission execution. The second area addresses SSA Operations, covering CRM aspects pertinent to monitoring and managing space objects and debris.

We recommend adopting a comprehensive approach to Crew Resource Management, and its underlying tenets. This approach aims to achieve a reduction in human errors, resulting in fewer mishaps and catastrophic outcomes due to improved identification and mitigation of risks. It will enhance team performance by fostering better leadership, communication, coordination, and decision-making among operators. Moreover, improved mission success rates and higher operational efficacy are expected outcomes through higher efficiency and effectiveness in mission execution.

This is accomplished by providing training that focuses on group dynamics including assertiveness, leadership, mission analysis, situational awareness, interpersonal communications, adaptability/flexibility, and decision-making.

6. Proposed CRM Training Framework for Space Operators

To ensure the success of space operations in increasingly complex environments, we recommend adopting a comprehensive approach to Crew Resource Management, and its underlying tenets. Training modules should include topics such as: Communication protocols (e.g., callouts, checkbacks); Leadership and followership; Conflict resolution and assertiveness; Decision-making under uncertainty; Cognitive workload and stress management; Cultural and interagency coordination; Each module should include simulated exercises, live rehearsals, and after-action reviews, supported by data collection for ongoing improvement.

Current Generation Applied: Error Management and Threat Models

In [37], The University of Texas Threat and Error Management Model (UT-TEMM) provides a valuable framework for classifying and managing operational errors. The model recognizes that:

1. Error is inevitable
2. Most errors stem from flawed communication, decision-making, or coordination
3. The key to safety is managing threats and mitigating errors, not eliminating them

CRM training built on this model trains individuals to identify latent threats, communicate proactively, and implement countermeasures such as briefings, cross-checks, and structured decision-making protocols. These are directly applicable to space missions, especially in multi-agency or coalition contexts.

Teaching Vigilance, Failure Detection, and Automation

Previously presented research in [9] and [13] show that passive monitoring of automated systems results in longer detection latencies and lower situational awareness. Operators are more effective at identifying anomalies when engaged in active control. This has direct implications for space operations where over-reliance on automation can degrade operator performance.

CRM combats this by enhancing vigilance through shared monitoring responsibilities, structured role assignments, and assertive communication. Training modules should address failure salience, automation complacency, and strategies for switching from passive to active control under stress.

Assuring Skill Retention in Low-Tempo Environments

Given that space domain operators often work in low-tempo environments punctuated by high-consequence events. This creates challenges for skill retention, especially for rarely used emergency procedures. Reference [10] shows that:

1. Skills degrade over time without rehearsal
2. Emergency and procedural skills degrade faster than routine skills
3. Overlearning and distributed practice improve long-term retention

Our proposed CRM training therefore incorporates regular scenario-based refreshers and cross-training to ensure readiness across varying retention intervals.

Organizational Culture and Leadership Integration

Organizational culture strongly influences CRM effectiveness. Safety culture must value openness, non-punitive error reporting, and continuous improvement. Leadership at all levels must reinforce CRM practices and integrate them into standard operating procedures, performance reviews, and mission planning [2].

Other courses of action

In addition to enhancing existing training with CRM, an opportunity exists to establish global CRM standards for space domain awareness functions. Developing these standards may reduce the risk of future incidents like the Iridium 33 collision. Enhanced collaboration between satellite operators and situational awareness enterprises is critical to avoid future mishaps.

These recommendations position CRM as a cornerstone for operational excellence in space, helping to mitigate risks and foster innovation in a highly dynamic environment where humans are inevitably part of every operational system.

7. Conclusion

Space operations are approaching a complexity threshold that mirrors other domains that rely on high performance teams in terms of error vulnerability and mission consequences. Integrating CRM into the training and operational doctrine of space domain personnel is no longer optional. Founded in aviation, extended to other domains, drawing on decades of experience, validated behavioral models, and human performance research, CRM offers a tested and adaptable framework to meet the evolving demands of space.

Therefore, training should include the concepts presented here, in existing training programs from other domains, and be informed by domain specific research that has not yet occurred.

An effective training syllabus [38]:

1. Embraces all operational personnel
2. Can be blended into all forms of training
3. Concentrates on attitudes and behaviors and impact on safety and mission ops
4. Use the crew/team as the unit of training

5. Requires active participation of all team members
6. Opportunity for operators to examine their own behavior, and to make decisions on how to improve operational teamwork
7. Addresses communications processes and decision behavior
 1. Briefings
 2. Inquiry/Advocacy/Assertion
 3. Crew Self-Critique (Decisions and Actions)
 4. Conflict Resolution
 5. Communications and Decision-making
8. Provides a common, repeatable framework for team building and maintenance
 1. Leadership/Followership/Concern for Task
 2. Interpersonal Relationships/Group Climate
 3. Workload Management and Situational Awareness
 4. Individual Factors/Stress Reduction

Implementation pathway

Given a large body of knowledge in behavioral psychology, we need to study domain operations to select and apply germane content. Study of space domain operators in their native professional environments to identify which tenets apply to their functional paradigms.

There are several academic institutions that feed the technical elements of space domain systems. At present there is no academic leadership in this discipline in space or any other domain. The University of Texas at Austin Human Factors Research Project is the last center of excellence that studied this specific subset of behavioral psychology. This center shut down in 2011 when its founder, Dr. Robert Helmreich retired. Unfortunately, he passed away in 2012. The baton is waiting to be picked up by another academic institution to carry the discipline forward at the pace of technological advancement.

Leaders in operational organizations must be willing to embrace and support the need for training about the human element. Therefore, effective implementation is most likely to prove effective when emphasis comes from all levels of an adopting organization.

Given existing precedents in other domains, a policy discussion to explore the potential for baselining non-government space domain operators is also warranted.

9. Definitions

Automation – A process that controls a function or task without human intervention [12].

Behavior – The way in which an animal or person acts in response to a particular situation or stimulus [1].

Behavioral Psychology – A systematic approach to understand the behavior of humans and other animals [39].

Cognitive Skill – Mental skills that are a prerequisite to performing a task or subtask. Cognitive skill has a substantial mental component [36].

Cognitive Load – The processing burden on the brain to make sense of information; the total amount of mental activity imposed on the working memory in any one instant [40].

Cognitive Task and Motor-induced State Fatigue – A psychophysiological condition that is characterized by a decrease in motor or cognitive performance and/or an increased perception of fatigue [41].

Common – "...[places and] resources available to groups of people to manage for individual and collective benefit" [42].

Complacency – "self-satisfaction, which may result in non-vigilance based on an unjustified assumption of satisfactory system state" [43].

Crew Resource Management – One way of addressing the challenge of optimizing the human/machine interface and accompanying interpersonal activities [43].

Distraction – A thing that prevents someone from giving full attention to something else [1].

Heuristic – A rule or piece of information used in or enabling problem-solving or decision-making [1].

Human Factors – The applied science that studies people working together in concert with machines [44].

Human Performance - The abilities, skills, and actions exhibited by individuals in various tasks or activities. It can be measured by factors such as accuracy, speed, productivity, and quality of work [45]. The ability to perform consistently at our best in the environment within which we are operating at that moment [46]. A series of behaviors executed to accomplish specific results (performance = behavior + results) [47].

Human Performance Curve (Yerkes-Dodson) – Depiction of the empirical relationship between stress and performance and that there is an optimal level of stress corresponding to an optimal level of performance [8].

Rule of Thumb – A broadly accurate guide or principle, based on experience or practice rather than theory [1].

Power Distance - The levels of authority and inequality within a hierarchical structure. It signifies the extent to which decision-making power is concentrated at the top and the acceptance of unequal rights between individuals of different positions. In high power distance cultures, individuals in lower positions have limited involvement in decision-making, and their fate is often determined by those in positions of authority [48].

Procedure – An established or official way of doing something [1].

Situational Awareness – The accurate perception and understanding of all factors and conditions in context of operations [49].

System – A set of things working together as parts of a mechanism or an interconnecting network [1].

Technique – Way of carrying out a task, especially the execution or performance of procedure[s] [1].

REFERENCES

- [1] Oxford Languages. (2025). Oxford English Dictionary Online. Retrieved from <https://languages.oup.com/>
- [2] University of Texas at Austin Human Factors Research Project: 235 Helmreich, R.L., Merritt, A.C., & Wilhelm, J.A. (1999). The evolution of Crew Resource Management training in commercial aviation. *International Journal of Aviation Psychology*, 9(1), 19-32.
- [3] Helmreich, R. L. (2000). *On Error Management: Lessons from Aviation*. *BMJ*, 320(7237), 781–785.
- [4] Helmreich, R. L., & Davies, J. M. (2004). *Culture, Threat, and Error: Lessons from Aviation*. *Canadian Journal of Anesthesia*, 51(6), R1–R4.
- [5] Gaba, D. M. (2000). *Anesthesiology as a Model for Patient Safety in Health Care*. *BMJ*, 320(7237), 785–788.
- [6] Sexton, J. B., Thomas, E. J., & Helmreich, R. L. (2000). *Error, Stress, and Teamwork in Medicine and Aviation: Cross Sectional Surveys*. *BMJ*, 320(7237), 745–749.
- [7] FAA. (2000). Order 9550.8 Human Factors Policy. Federal Aviation Administration.
- [8] Nickerson, R. (2023). The Yerkes-Dodson Law and Modern Work Environments. *Cognitive Science Review*.
- [9] Gawron, V. J. (2019). *The Watching—A Review of the Vigilance Research*. The MITRE Corporation.
- [10] Gawron, V. J. (2019). *Lessons Lost: The Doing—A Review of Skill Retention Research*. The MITRE Corporation.
- [11] Gawron, V. J. (2019). *Lessons Lost: Is Something Wrong? Failure Detection Research*. The MITRE Corporation.
- [12] Billings, C. E. (1991). *Human-Centered Aircraft Automation: A Concept and Guidelines*. NASA Technical Memorandum 103885. National Aeronautics and Space Administration.
<https://ntrs.nasa.gov/citations/19910015714>
- [13] Gawron, V. J. (2019). *Automation in Aviation—Definition of Automation*. The MITRE Corporation
- [14] Wiener, E. L. (1989). *Human Factors of Advanced Technology (“Glass Cockpit”) Transport Aircraft*. NASA Contractor Report 177528.
- [15] Endsley, M. R., & Kiris, E. O. (1995). *The Out-of-the-Loop Performance Problem and Level of Control in Automation*. *Human Factors*, 37(2), 381–394.
- [16] Gawron, V. J. (2019). *Automation in Aviation—Accident Analyses*. The MITRE Corporation.
- [17] FAA (June 1996). FAA Human Factors Team Report on interfaces between flight crews and modern flight deck systems. Washington, DC.
- [18] Funk, K. & Lyall, B. (1998). Human factors issues of flight deck automation, (pp. 1, E23/1–8). Proceedings of the 17th Digital Avionics Systems Conference.
- [19] Sheppard, R. (2023). Subsequent Assessment of the Collision between Iridium 33 and COSMOS 2251. Proceedings of the 2023 Advanced Maui Optical and Space Surveillance Technologies Conference.
- [20] Gawron, V. J. (2019). *Automation in Aviation—Guidelines*. The MITRE Corporation.
- [21] Feggetter, (1982). A. J. A method for investigating human factor aspects of aircraft accidents and incidents. *Ergonomics*, 25(11), 1065–1075.

- [22] NASA. (1968). *Report of Apollo 204 Review Board to the Administrator, National Aeronautics and Space Administration* (NASA Historical Reference Collection No. SP-323). National Aeronautics and Space Administration.
- [23] Levine, A. (2016). *Apollo 1: The tragedy that put NASA on a path to success*. HarperCollins.
- [24] Long, N. (2007). *A managerial approach to NASA's cultural changes open-system model* (Doctoral dissertation, Monterey, California. Naval Postgraduate School).
- [25] NASA. (1970). *Apollo 13 Review Board Report*. NASA Historical Reference Collection.
- [26] Vaughan, D. (1997). *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. University of Chicago Press.
- [27] Rogers Commission. (1986). *Report of the Presidential Commission on the Space Shuttle Challenger Accident*. Government Printing Office.
- [28] Columbia Accident Investigation Board. (2003). *Columbia Accident Investigation Board Report, Volume I*. NASA and Government Printing Office.
- [29] Kelso, T. S. (2010). Analysis of the Iridium 33–Cosmos 2251 collision. *Advances in Space Research*, 45(10), 1409-1417. <https://doi.org/10.1016/j.asr.2010.01.008>
- [30] Cummings, M. L. (2014). *Artificial intelligence and the future of air traffic control*. In K. D. J. Gill & S. H. Barnes (Eds.), *Systems Engineering and AI in Aviation* (pp. 47-66). Wiley.
- [31] Cummings, M. L. (2014). Automation bias in intelligent time-critical decision support systems. AIAA Infotech@Aerospace Conference.
- [32] Cummings, M. L. (2014). Man versus machine or man *with* machine? *IEEE Intelligent Systems*, 29(5), 62–69. <https://doi.org/10.1109/MIS.2014.82>
- [33] Weeden, B., & Chow, T. (2012). Developing norms of behavior for space activities: Progress, challenges, and next steps. *Astropolitics*, 10(1), 44–59. <https://doi.org/10.1080/14777622.2012.647012>
- [34] Mashiku, A., Highsmith, D., Small, J. (2024). NASA CARA Human Factors. Training Materials.
- [35] Flight Safety Foundation. (2000). ALAR Briefing Note 2.4: Interruptions/Distractions.
- [36] Federal Aviation Administration. (1998). *Developing Advanced Crew Resource Management (ACRM) Training Manual*.
- [37] NASA/University of Texas Human Factors Research Project. University of Texas Threat and Error Management Model (UT-TEMM). Retrieved from <https://www.psy.utexas.edu/psy/helmreich/nasaut.htm>
- [38] Donnelly, A. M. (2024) On-the-loop: Defining the Requirements to Coexist. Unpublished
- [39] Wikipedia. (2025). Behavioral Psychology. Retrieved from <https://en.wikipedia.org/wiki/Behaviorism>
- [40] Cognition Today. (2022). What is Cognitive Load and How to Reduce It. Retrieved from <https://cognitiontoday.com>
- [41] Behrens, M., et al. (2022). Fatigue and Cognitive Performance in High-Stakes Environments. *Journal of Neuroergonomics*, 5(1).

- [42] Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.
- [43] NASA. (2017). *Human Factors Considerations in the Design and Evaluation of Flight Deck Displays and Controls*. NASA/SP-2017.
- [44] Federal Aviation Administration (2004). *Advisory Circular (AC) 120-51E: Crew Resource Management Training*.
- [45] Fiveable. (2025). *Human Performance Overview*. Retrieved from <https://fiveable.me>
- [46] Kelly, J. (2023). *Performing Under Pressure: The Human Performance Curve*. *Psychology of Excellence Series*.
- [47] U.S. Department of Energy. (2009). *Human Performance Improvement Handbook Volume I: Concepts and Principles*.
- [48] Science Direct (2025). *Power Distance*. Retrieved from <https://www.sciencedirect.com/topics/social-sciences/power-distance>
- [49] Avibary.com. (2025). *Situational Awareness in Aviation*. Retrieved from <https://avibary.com>
- [50] Maurino, D., Reason, J., Johnston, N., & Lee, R. (1995). *Beyond Aviation Human Factors*. Aldershot: Ashgate Publishing.
- [51] Sexton, J. B., & Helmreich, R. L. (2000). *Analyzing Cockpit Communications: The Links Between Language, Performance, Error, and Workload*. *Journal of Human Performance in Extreme Environments*, 5(1), Article 6.