Evidence Report:

Risk of Performance Errors Due to Training Deficiencies

Immanuel Barshi, Ph. D   *NASA Ames Research Center*

Human Research Program

Space Human Factors and Habitability Element

Approved for Public Release: <mark>Month DD, YYYY</mark>

TABLE OF CONTENTS

Risk Title

The ***Risk of Performance Errors Due to Training Deficiencies*** is identified by the National Aeronautics and Space Administration (NASA) Human Research Program (HRP) as a recognized risk to human health and performance in space. The HRP Program Requirements Document (PRD) defines these risks. This Evidence Report provides a summary of the evidence that has been used to identify and characterize this risk.

Risk Statement

Given that training content, timing, intervals, and delivery methods must support crew task performance, and given that training paradigms will be different for long-duration missions with increased crew autonomy, there is a risk that operators will lack the skills or knowledge necessary to complete critical tasks, resulting in flight and ground crew errors and inefficiencies, failed mission and program objectives, and an increase in crew injuries.

Risk Overview

Human error has been implicated as a causal factor in nearly two thirds of mishaps across NASA (Chandler, 2007), and similar situations exist in related domains like commercial and military aviation (70-80% of incidents and accidents involve human error directly, and 100% of accidents and incidents involve human limitations in some way, see, e.g., Maurino, Reason, Johnson, & Lee, 1995). In a significant proportion of incidents involving human error, incorrect procedure execution played a role.

Procedure execution errors (both of omission and commission) result from some combination of: inadequately designed tasks; inadequately designed procedures or tools; incomplete, inaccurate, or difficult-to-use documentation; fatigue, stress, injury, or illness; insufficient training (including lack of training for unanticipated operations); degradation of trained skills or knowledge; or inadequate understanding of the operational environment.

Historically, spaceflight operations have mitigated some of these effects in at least two ways: specially-trained crewmembers are assigned to missions and/or rotated into the operational environment when complex, mission-critical tasks must be performed; and, execution of such procedures is closely monitored and supported by flight controllers on the ground who have access to a broader and deeper pool of information and expertise than any individual operator.

However, emerging mission architectures include long-duration operations in deep space. Such operations do not allow for assignment of new crewmembers or even for the development and validation of new training on the ground. Further, delays in communication will have a disruptive effect on the ability of earth-based flight controllers to monitor and support space operations in real time. As a result, it is necessary to develop an understanding of how training can be tailored to better support long-duration deep-space operations (including the extent to which materials, procedures, and schedules of training should be changed from current practices). Note, long-duration missions in low-earth orbit (LEO), such as operations on the International Space Station (ISS), and even lunar operations are of a different nature than deep-space operations given the large difference in communication delays and the ability to return crewmembers to Earth within a very short time.

The evidence presented in this chapter describes issues related to the risk of performance errors due to training deficiencies. Contributing factors regarding training deficiencies may pertain to organizational process and training programs for spaceflight, such as when training programs are inadequate or unavailable. Furthermore, matching between task and learning and memory abilities is a contributing factor. A mismatch can occur when the individual's knowledge acquisition strategies and predispositions are inconsistent with mission demands. Learning, namely the relative efficiency with which an individual acquires new information and with which the relatively permanent adjustments in behavior or thinking are made, may not be consistent with mission demands. Thus, if training deficiencies are present, the likelihood of errors or of the inability to successfully complete a task increases. Importantly, there is an increase in the overall risk to the crew, the vehicle, and the mission.

Risk is always a function of likelihood and consequences. Two important considerations in assessing the likelihood of the Training Risk are mission duration and the extent to which the crew must be autonomous (due to communications lag). The likelihood increases with mission duration due to increased intervals between training and operation where no refresher or just-in-time training is in place; also, unexpected or emergent situations for which no applicable training exists are more likely in a long-duration mission. Likelihood also increases with higher levels of crew autonomy (due to less opportunity to obtain guidance/refresher training from ground support). When both of these factors are present (long-duration missions that require high crew autonomy), the likelihood of an adverse outcome is at its maximum. Accordingly, the Asteroid (Near Earth Asteroid (NEA)/Near Earth Object (NEO)) mission architecture (12+ month duration; up to 30 second communication delay) and the Mars mission architecture (3 year duration; up to 22 minute communication delay) are more likely to produce adverse outcomes. Lunar and ISS missions may have moderate to long-durations, but no appreciable communication delay; therefore, crew autonomy may not be as critical an issue.

On the consequence side of the risk function, for missions where communication with ground support is fast and reliable and where rotation of crew into operations is possible, training can be made more effective (in terms of length of time required for the acquisition of new knowledge and skills, durability of the of the acquired skills, and the ability to transfer learned knowledge and skills to new situations). Furthermore, it's possible to compensate for inadequate training utilizing current training practices: Ground support personnel can monitor operations in real-time and offer guidance, instruction, or other information. Freshly-trained crew can be provided from the ground to execute sensitive and/or complex operations, and those operations can be executed in close temporal proximity to extensive ground-based training. However, when crews are remote from Earth and communication involves delays or is unreliable, those augmentations and mitigations to training are not possible. If inadequate crew training results in tasks being executed incompletely, incorrectly, or inefficiently, operations will be hindered. Only delayed ground support will be available to assist with development of workarounds or to provide guidance, instruction, or other information required to correctly complete a deferred task, by which time it might be too late.

Long-duration deep-space missions also make the attenuation of adverse outcomes (particularly, re-supply or rescue operations) difficult or impossible, so the severity of outcomes may be greater for such operations. In general though, training-related errors are highly likely to cause time losses or inefficiencies and are less likely (but able) to cause serious damage to vehicles, habitats, or other equipment. In extreme cases (and, as is often the case, in combination with other factors), such errors could have significant impacts on mission objectives and crew safety.

When other factors do not contribute to a situation to create high-severity consequences (e.g., a skipped step might have dire consequences in an emergency procedure where time is critical), most training inadequacies will lead to moderate impacts to operations (e.g., extended task times and the need for additional help from other crew or ground). Note that consequences become potentially far more serious in dynamic flight phases such as launch, docking, and landing, when there is very little time available to correct mistakes. It is also possible that a misstep during a medical operation would have personnel and health consequences; such consequences would be hard to predict.

Levels of Evidence

The levels of evidence presented in this chapter are based on the Levels of Evidence in the NASA Risk Management and Analysis Tool. These are: Case Study, Expert Opinion, Terrestrial Data, Expert Data, and Spaceflight Incidence. Evidence presented in this chapter encompasses lessons learned from 50 years of spaceflight experience and ground-based research related to the risks due to performance errors resulting from training deficiencies. Portions of the evidence

consist of summaries of subjective experience data, as well as non-experimental observations or comparative, correlation, and case or case-series studies. It should be noted that some evidence in this chapter is derived from the Flight Crew Integration (FCI) ISS Life Sciences Crew Comments Database. Although summaries of ISS crew feedback are presented as evidence, the database is protected and is not publicly available, due to the sensitive nature of the raw crew data it contains.

Levels of Evidence may include, but are not limited to:

- Case study
- Expert data
- Expert opinion
- Spaceflight incident data
- Terrestrial data
- Modeling

Evidence

Training is the systematic acquisition of knowledge, skills, and attitudes. It is about providing those being trained with the right competencies necessary to successfully perform a specific job or task (Salas, Wilson, Priest, & Guthrie, 2006). When the right knowledge is not transferred during training, errors occur.

As early as 1953, in the early days of human factors research, training had already been recognized as a critical issue. In his seminal work on human factors in air transportation, McFarland (1953) dedicates a whole chapter to training issues. He notes that "little attention has been given to the human factors involving an airman's understanding of his environment and the physical factors influencing his efficiency" (p. 152). More recent works continue to echo the same concerns. Training deficiencies are clearly pointed out in the analysis of accidents and incidents in aviation (e.g., Barshi & Loukopoulos, 2012; Dismukes, Berman, & Loukopoulos, 2007; Loukopoulos, Dismukes, & Barshi, 2009a, 2009b), other modes of transportation (National Transportation Safety Board), and in other high-risk industries (e.g., Grote, 2009; Helmreich, & Merrit, 1998; Reason, 1997).

Most accidents and incidents involve some loss of situation awareness (SA). Such loss of SA is often cited as causal in aviation accidents as well as in Mission Control Center errors. Analyses of general aviation accidents that have been caused by the pilot's lack of SA cited inadequate general aviation training that did not effectively address how to improve SA. Analysis was conducted to determine the nature of the problems that occur in situations where SA was lost, as

well as the differences in SA between pilots who perform well and those who do not. The results yielded four key recommendations for how SA could be improved through training. The first recommendation is to provide training that will allow the pilots to develop good task-management strategies to deal with interruptions, distractions, and overall workload that can pose a high threat to SA. The second recommendation is development of comprehension: providing pilots with the tools necessary to properly gauge the temporal aspects of the situation, the risk levels involved and personal and system capabilities for dealing with the situation. The third recommendation is to provide pilots with skills to project and plan, which will allow them to actively seek important information in advance of a known immediate need for it, and plan for contingencies. The final recommendation is to encourage information seeking and self-checking activities, i.e., skills that will help pilots notice trends and react to events more quickly. Development and implementation of training programs that focus on these four recommendations should effectively improve SA in pilots (Endsley & Garland, 2000).

Another specific area in which training deficiencies have been identified in accident and incident analysis is team communication. It has been reported that deficient communication between the ground and crew can cause frustration and negatively affect performance. This can be due to ground operators having difficulty identifying information related to task duration, which in turn frustrates the crew and ground personnel because the perception of task duration is different between those developing timelines and those executing the task. Many times crewmembers have not been able to identify information regarding what the ground could assist with and what tasks could be automated to facilitate crew productivity (Rando, Baggerman, & Duvall, 2005). The most efficient and effective teams (e.g., aviation, military, design teams) manage to coordinate their activities with just enough, but not too much communication (Entin & Serfaty, 1999; Orasanu & Fischer, 1992; Patrashkova-Volzdoska, McComb, Green, & Compton, 2003). Thus, proper training in communication is crucial, because a team is at risk in the absence of task-specific procedures and training for assuring appropriate communication, especially in the case of distributed teams in which shared context cannot carry some of the communication function. To overcome this risk, task procedures, training, and tools must be designed in such a way that essential information is communicated while keeping down the process cost, thereby minimizing the added workload required to communicate among team members.

In a study by Schaafstal, Schraagen, and van Berlo (2000), Navy weapons engineers' troubleshooting problems were attributed to:

- Insufficient training,
- Technical documentation written from an engineering viewpoint instead of a maintenance viewpoint, and
- A gap between the theoretical knowledge and application of this knowledge in a real-life situation. That gap is a clear training deficiency.

Much of the research literature on training is motivated by current training practices aimed at current training needs. There is no experience with long-duration space missions where crews must practice semi-autonomous operations, where ground support must accommodate significant communication delays, and where so little is known about the environment. Thus, not only must we resolve known deficiencies in the current state-of-the-art in training, we must also develop robust methodologies and tools to prepare our crews for the unknown. The research necessary to support such an endeavor does not currently exist, but existing research does reveal general challenges that are relevant to long-duration, high-autonomy missions.

For example, learning is highly specific to the conditions under which it occurred, especially when the learning involves procedural, as opposed to declarative, information. To account for this specificity, Healy and Bourne (1995a; see also Healy, 2007; Healy, Wohldmann, & Bourne, 2005) proposed a procedural reinstatement principle, according to which training on one skill does not transfer to another related skill unless the procedures required by the two skills overlap. This principle is clearly related to other principles and theories in the literature, including: Thorndike's (1906) theory of identical elements (see also Rickard & Bourne, 1995, 1996; Singley & Anderson, 1989); Tulving and Thomson's (1973) encoding specificity principle; Morris, Bransford, and Franks's (1977) transfer-appropriate processing principle (see also McDaniel, Friedman, & Bourne, 1978; Roediger, Weldon, & Challis, 1989); Proteau, Marteniuk, and Lévesque's (1992) specificity of practice theory involving sensorimotor representations for motor learning; and Kolers and Roediger's (1984) theory involving procedures of the mind.

Performance also changes with level of stress on the trainee. At low levels of stress, performance might be poor, but as stress increases gradually, performance improves. At a certain point, stress level is optimal for performance on a given task. Beyond the optimum, additional stress might degrade performance, and when stress becomes extreme, the trainee might "choke" or panic (Staal, Bolton, Yaroush, & Bourne, 2008). However, stress has been shown to affect speed and accuracy of responses differently. For example, the stress that comes from fatigue developed as a result of continuous work on a task leads to faster but less accurate performance (a speed-accuracy tradeoff; see Healy, Kole, Buck-Gengler, & Bourne, 2004). Similarly, Wolfe, Horowitz, Cade, and Czeisler (2000) found that sleep deprivation led to an increase in errors on a visual search task for a target among varying numbers of distractors as well as to a reduction in the slope of the function relating response time to the number of distractors (see also Horowitz, Cade, Wolfe, & Czeisler, 2003). Thus, sleepy observers responded quickly but carelessly. Consequently, adding stressors to a task could be harmful (e.g., in the case of accuracy) or beneficial (e.g., when speed is the primary requirement) depending on what aspects of the task are most crucial and on the ambient level of stress. Stress and sleep loss contributing factors are discussed in the evidence book for the HRP risk on sleep, thus is not addressed here.

Further, declines in performance are associated with increasing intervals between training and performance. Multiple causal factors contributed to Mir-Progress collision (Ellis, 2000). The Russian spacecraft Progress 234 collided with the Mir space station, causing the pressure hull to rupture, and nearly causing the Mir to be abandoned (Ellis, 2000; Shayler, 2000). One of the causal factors of this event resulting in a near-catastrophe was the training-performance interval. The crew last received formal training four months before the docking incident. Such skill degradation is likely to be exacerbated as mission duration increases, and crews place greater reliance on onboard automation.

Performance declines, exhibited by increased response time (or decreased accuracy), have been known since the time of Ebbinghaus (1885/1913), who used a measure of savings (i.e., the amount of relearning required to achieve the criterion level of performance during original learning). Subsequently, this relationship between response time and retention interval was described as a power law (Wickelgren, 1974), $R = d + fT\text{-}g$, where R is response time, T is the retention interval, d is the criterion of original learning, f is a scaling parameter, and g is the rate of forgetting. This Power Law of Forgetting (Wixted & Carpenter, 2007; see also Rubin & Wenzel, 1996) can be thought of as the inverse of the Power Law of Practice (Newell & Rosenbloom, 1981), which describes the acquisition process for most skills (the relationship between trials of practice and time to make a correct response is a power function, $R = aN\text{-}b$, where R is response time on trial N, a is response time on trial 1, and b is the rate of change). The understanding of these relationships is crucial for the design of effective training.

Current shuttle astronaut training is largely based on extended practice (e.g., http://www.spaceflight.nasa.gov/shuttle/support/training/ascan/2004/index.html; personal communications). Shuttle mission success was partly due to a long lead-time (2-3 years) for training, short (1-2 weeks) and highly scripted Shuttle missions, and continuous communication with ground support. For ISS missions, most of the training is completed on the ground before the mission. Moreover, pre-launch training is mostly task-specific rehearsal, often distributed over time and place, which minimizes coherence and continuity.

Concerns about training issues continue to be a common complaint in debriefings of crews returning from space missions. For example, interviews with crewmembers who have served as Crew Medical Officer (CMO) indicate that following initial training, there are no formal reviews of the entire course if/when they are assigned to another mission, and that there is no formal assessment of the effectiveness of current CMO training and onboard refresher training modules for long-term retention of space medical training. It's possible that such deficiencies are the result of underestimating the importance of training due to the history of low frequency of medical events in space, and a certain comfort level due to ground medical expertise always being available when needed. However, that will not be the case with future, long-duration space missions. Similarly, minimal to no influence of current understanding of human learning, skill

acquisition and retention can be seen in recent "just-in-time" training for new experiments sent to ISS. For example, refresher training for the Dust and Aerosol Measurement Feasibility Test (DAFT) experiment was uplinked to ISS on February 3rd, 2005; it consisted of only three PowerPoint slides, each with a few bullet points, one slide with two pictures of relevant equipment, and one slide of the general process flow chart. Multiple challenges arose, leading to the premature termination of the protocol (see, e.g., Evans et al. 2009; Urban et al. 2005). Such training will not be adequate to support long-duration missions where many potential tasks cannot be predicted and trained in advance, and where ground support may not be immediately available. Furthermore, given the stress and problems associated with current training practices for the individual crewmembers and their families[1]; shorter, more efficient, and more effective training is needed to better support our crews. Research and development is required to maximize training coherence for effective retention and transfer. Furthermore, skill-based training must replace task-based rehearsal to support generalizability and the ability to deal with unexpected, untrained-for emerging tasks and opportunities. Crews must develop appropriate responses and procedures as situations demand, drawing upon more in-depth understanding of spacecraft systems and operations. More extensive cross training will also enhance situational resiliency.

The Training Continuum

Training can be seen as distributed along a continuum, from initial pre-mission training, through ground and in flight refreshers, to onboard initial training, to just-in-time training, and all the way to performance support tools. For optimal results, careful distribution of training topics and training methodologies is required across the full continuum of training opportunities. The relevant issues at each point along this continuum are different. For the initial, pre-mission training, key questions have to do with task-based training vs. skill-based training, with optimal use of simulation facilities and the methodology by which the right level of simulation fidelity can be selected given the operational context and the individual learner, as well as the right training methodology and delivery system (e.g., classroom vs. computer based training (CBT)). Onboard initial and refresher training will be significantly constrained by onboard delivery technologies, though virtual environments hold great promise. However, given the limited

---

[1] *The current training regime is very stressful for crewmembers and their families as it involves many years of intensive and intense work, frequent overseas travel, and long absences from home. Crewmembers often comment that the missions themselves are far easier on everybody than the years of training.*

ground-based experience that has been gained thus far with such systems, key questions about training methodologies and delivery mechanisms remain, as well as about differences in acquisition and retention between on-ground and onboard training sessions. For refresher training in particular, questions about which topic should be refreshed and at what interval require a systematic methodology for determination. Just-in-time training requires the ability to expect the unexpected. That is, training can only be developed for expected tasks and situations. Just-in-time training will be needed for low likelihood events and situations that are not time critical. Because not all such events can be anticipated in advance, methods for the crew to develop their own training for such occasions must be developed for cases when communication delays prevent the up-link of such training from the ground. Somewhat similar issues exist for the development of performance support tools, either for situations that don't justify training, or for when time is critical and the crew does not have the opportunity to receive training prior to operations. Again, determining which events can be handled using performance support tools and how such tools can best be designed requires systematic methodologies that do not yet exist.

The risk of performance errors due to training deficiencies includes two primary contributing factors: 1) Organizational Training Issues/Programs; and 2) Matching Between Tasks and Learning and Memory Abilities. The contributing factors were derived from the Department of Defense (DoD) Human Factors Analysis and Classification System (HFACS), the industry standard for human error categorization. (Department of Defense, 2005). The evidence about risk reduction presented in this report is organized around two types of causal risk factors, selected from the HFACS categories of error (Shappell & Wiegmann, 2000). This classification system attempts to identify the point or points in a causal chain of events that produced an accident, typically with behavior identified as an error after the fact. This approach focuses on explaining events after they happen, and providing a causal chain in this explanation.

Contributing Factor 1: Organizational Training Issues/Programs

Organizational training issues/programs is a factor when training programs are inadequate or unavailable. Training that aims at the acquisition of durable skills often leads to high specificity (see, e.g., Healy and Bourne, 1995; Vogel and Thompson, 1995). Thus, training for tasks that will be performed under extreme conditions and environments can be less than adequate for other mission objectives. This is due to the fact that simulated environments and ground-based full-scale models or mockups cannot be completely representative of flight conditions. Representing a true 0g environment on the ground has presented many challenges for training, so current simulations facilities and methods may not be adequate for preflight training. Figures 1 and 2 depict training both in ground based mockups and trained operations being performed onboard the ISS.

Figure 1. JSC2007-E-46438 — Astronaut Peggy A. Whitson, Expedition 16 commander, and cosmonaut Yuri I. Malenchenko, flight engineer representing the Russian Federal Space Agency, participate in a ground based training session with the onboard communication system in a Russian module mockup at the Gagarin Cosmonaut Training Center, Star City, Russia (NASA Human Spaceflight Gallery, 2008).

Figure 2. ISS016-E-022130 — Cosmonaut Yuri I. Malenchenko, Expedition 16 flight engineer representing the Russian Federal Space Agency, performs previously trained, real time operations onboard ISS with the onboard communication system while working in the Russian Zvezda service module of the ISS (NASA Human Spaceflight Gallery, 2008).

Stowage problems on ISS are one example of 1-G training failing to build adequate skills needed for task execution in 0-G. As documented in the FCI ISS Life Sciences Crew Comments Database, a true representation of the stowage of equipment and materials onboard the ISS is very difficult to achieve on the ground and can create issues for the crew. Stowage mockups in 1g are limited because gravity restricts operations, translation, and stowage placement in the training facilities. Given the constraints of a 1G-based translation path, it is not possible or safe to place things where they would potentially be stowed onboard the ISS. On-orbit there is the benefit of weightlessness, which allows stowage of items on any axis with proper restraint. The crew can translate through the available volume and position their bodies to move around obstructions or protrusions in the translation paths. Additionally, while on-orbit, some of the

stowage lockers are packed tightly, making it difficult to re-stow items due to the lack of gravity working against the crewmember. Similarly, crews often have trouble with items floating off during retrieval or re-stow. In a 1G environment, stowage does not behave the same as in a 0G environment. However, tasks and procedures are written based on what is known from testing in a 1G environment, and so is the training; unfortunately, a mock-up may not be the best representation of what it will be like for the crewmembers while on-orbit. Given the gravitational differences between Earth and orbit, and disconnects between ground training and actual life on-orbit, the crewmembers have a lot to learn once onboard the ISS. The result is that, upon arrival on ISS, the crew often has difficulty managing stowage and operating nominally, leading to loss of precious time, inefficiencies, and crew frustration.

Long periods of disuse lead to skill decay (see, e.g., Winfred, et al. 1998). The passage of time and the lack of opportunity to rehearse or refresh acquired knowledge or skills will result in performance decrements due to forgetting what was learned. Training programs that do not account for degradation of learned skills or knowledge (e.g., by including refresher training or by providing just-in-time training rather than advanced training on the ground) may result in inferior task performance. In addition, fragile memory structures crumble under stress and fatigue (see, e.g., Staal, 2004). Thus, training for long-duration missions must be robust, and must support extensive memory structures that can withstand the effects of fatigue and stress likely to be experienced during such missions. Because rote learning leads to fragile memory structures, training for long-duration missions will have to take a different approach. Thus, onboard refresher training as well as just-in-time training and performance support tools must be developed to support long-duration missions. Research must be conducted to determine the sensitivity of different skills to disuse, and the refresher schedule required. Furthermore, to support long-duration missions we must develop an understanding of the kinds of tasks that must be trained in advance, those that can be trained enroute, and those that can be trained just-in-time or even supported while being performed without prior training.

More generally, whether or not the correct training methods, materials, and platforms are available, qualified instructors and valid evaluation methods are also needed to ensure adequate training. Given that methods, materials, and platforms are still in development for long-duration mission concepts, it may be problematic to find or train instructors in order to promptly and properly support training programs for such future missions. However, the benefits of developing and validating systematic approaches to training skills and transfer of skill between situations far outweigh the potential costs and consequences of not being able to perform tasks on missions with high crew autonomy. Furthermore, developing guidelines for adequate distribution of training topics across the full continuum of training opportunities (from initial pre-mission training, through ground and in flight refreshers, to onboard initial training, to just-in-time training, and all the way to performance support tools) will substantially increase training efficiency and effectiveness. Determining the proper methodologies to support such distribution

will mitigate further risk posed by current practices. Assuring these benefits of new training approaches requires research.

Although training issues are covered in crew debriefs, systematic data on effectiveness of training are not collected or analyzed2. Thus, performance in spaceflight is not specifically assessed to validate current training practices. Skill-based training approaches have not been researched, and schedules for refresher training are not established; adaptive training technologies and delivery systems for onboard training have not been developed. Moreover, optimal strategies for training problem solving and decision making in space operations have not been studied. However, studies done by the HRP Space Human Factors and Habitability (SHFH) Space Human Factors Engineering (SHFE) Training Task to support the development of Flight Controller training in problem solving and decision making demonstrate significant promise (see, e.g., Schmidt, et al. 2011).

Task-based training does not necessarily lead to the development of generalizable skills (see, e.g., Healy and Bourne, 1995). Thus, training that is based on specific tasks might fail crewmembers faced with unexpected new emerging tasks. Because not all tasks on a long-duration mission can be anticipated, training should focus instead on generalizable skills. Although we know that this needs to be done, we don't yet know how to do it. Research is required to develop appropriate generalizable skill-based training.

A further complication arises as a result of the novel technologies and operational scenarios that will exist for deep-space missions. Extensive automated and robotic systems may be used to assist crew in their work. Given that these technologies are still emerging, it is likely that modeling and simulation platforms that can be leveraged for training on such systems will be limited or lacking. Even well-designed systems that address the right problem can produce accidents if humans do not understand what the automation/robot is doing or how control is distributed between human and automation/robot. Even experienced, skilled, and motivated users might make an error if training does not provide a complete and accurate model of the automation and human-automation integration, or if direct "hands-on" training with the system (or a simulation of the system) is not available. Training must align the user's mental model of the automated or robotic operations with how the automation/robot is designed to function (see, e.g., Billings, 1997). This includes teaching specific procedures, instilling in trainees an

---

[2] *Furthermore, crew performance during training is not always measured objectively, and training records are extremely limited. So assessing training effectiveness is a serious challenge.*

understanding of the environment and work to be done, and communicating a deep functional understanding of the automated systems. Training should also include information about how control is distributed between user and automation, how the user may change control, what other factors influence control state, and how to determine the current state of control. Some applications may require training for supervisors of automation or robotic behavior, not just operators.

Clearly, more research is necessary to disentangle the various relevant issues so that specific training recommendations can be used to guide organizational training programs.

Contributing Factor 2: Matching Between Task and Learning and Memory Abilities

Matching between task and learning and memory is a factor when the individual's relative efficiency with which new information is acquired and relatively permanent adjustments made in behavior or thinking, are not consistent with mission demands. Task design is often driven by the constraints of the technology involved. When the constraints of the human operator and those of the operational environment are not taken fully into account, problems arise. Furthermore, when training is designed strictly to meet performance criteria at the end of training, long-term retention is compromised (Healy and Bourne, 1995). Thus, when training is focused on a task that is technologically driven and is only aimed to "pass the class," trainees' ability to perform adequately following a long retention interval is greatly reduced increasing the risk of error and of compromising mission objectives.

Research has shown that training aimed at long-term retention often results in high specificity, and thus low generalizability (e.g., Vogel and Thompson, 1995). Specificity might be desired when all tasks are well known and understood in advance; that, however, cannot be the case for future long-duration space missions. If anything, we know that our crews will face situations and tasks we do not currently have the tools or the ability to foresee. As a result, future training should focus on generalizability of skills, rather than specificity. Current research has made important, albeit small, steps towards understanding generalizability, but we are still far from having a clear methodology for effective training of generalizable skills (e.g., Healy and Bourne, 2012).

One small example of the issues associated with this contributing factor has to do with the differences in the cognitive mechanisms underlying recognition and recall. The difference between recognition and recall was identified in the early days of psychological research (e.g., MacDougal, 1904); studies have shown that it is much easier to recognize than to recall. As MacDougal (1904) describes: "the name which cannot, by the greatest racking of memory, be brought back to consciousness is thus uttered spontaneously and without hesitation when the bearer is again met face to face" (p. 229). Similarly, when a task involves the manipulation of an

interface and contains a sequence of several steps, an operator can master that sequence through practice during training, by recognizing the correct buttons to push or switches to set. However, following a retention interval of disuse, as would be expected with many tasks during a long-duration space mission, the operator now has to recall those steps in the correct sequence, and no longer has the advantage of simply recognizing them. The memory structures constructed during training are thus critical to performance. Such memory structures can be very fragile and vulnerable to forgetting (see, e.g., Barshi & Loukopoulos, 2012; Loukopoulos, Dismukes, & Barshi, 2009a), even when the operator is an expert performer with extensive experience in the domain (see, e.g., Dismukes, Berman, & Loukopoulos, 2007). Hence, training that isn't sensitive to the potential mismatch between task demands and the operator's learning and memory constraints can lead to increases in the risk of error and of compromising mission objectives.

Computer-Based Modeling and Simulation

Understanding and predicting human-system performance and identifying risks that may be inherent in a concept or a design is often achieved via computer-based modeling or simulation. The use of human performance models can result in significant lifecycle cost savings as compared to repeated human-in-the-loop evaluations, but accurately modeling the human is extremely difficult. In the SHFE domain, modeling and human-in-the-loop evaluations must be used in concert. We do not have high-fidelity human performance models, and most of those existing models have not been sufficiently validated or certified. Accordingly, models must be used in a limited fashion – i.e., to help determine the critical areas that should be addressed through the more costly, but more representative human-in-the-loop evaluations. As mentioned above, modeling and simulation platforms can be leveraged for training on emerging technologies that are still in development.

Risk in Context of Exploration Mission Operational Scenarios

Future exploration-mission scenarios will increase in duration and in distance from earth. This will require developing new technology, new work methods, and new ways of ensuring that these novel elements are suitably integrated. This absolutely includes the development and application of proper training methodologies both on the ground in preparation for missions and while in space. Missions carried out in space will need greater flexibility and less dependence on ground support, and new interaction between ground-based resources and crew will also be needed and require increased pre-flight and onboard training. Risks from inadequate design of human-technology interaction will increase as mission requirements become more demanding and as missions are carried out in unfamiliar circumstances substantially different from our experience base. Human factors principles will need to be extended and applied to reduce risk.

Gaps

Potential gaps related to training include, but are not limited to:

- Lack of modeling and simulation platforms that can be leveraged for training on emerging technology and systems
- Inadequate or unavailable training programs
- Inconsistencies with training, individual attributes and mission demands

Training gaps will need to be coordinated with current and emergent gaps in other SHFE projects (e.g., Risk of Inadequate Design of Human and Automation/Robotic Integration and Risk of Inadequate Human-Computer Interaction) and HRP elements (e.g., Behavioral Health and Performance's Risk of Performance Decrements Due to Inadequate Cooperation, Coordination, Communication, and Psychosocial Adaptation within a Team and Exploration Medicine and Exploration Medical Capability). As NASA migrates toward longer duration missions beyond low-earth orbit, challenges associated with communication latencies and skill degradation will require cross-disciplinary solutions.

A summary of all SHFE gaps can be found in the Human Research Roadmap Content Management System at http://sa.jsc.nasa.gov/hrrcms/.

Conclusion

The evidence presented in this chapter describes issues related to the risk of performance errors due to training deficiencies. Contributing factors regarding training deficiencies may pertain to organizational process and training programs for spaceflight, such as when training programs are inadequate or unavailable. Furthermore, matching between task and learning and memory abilities is a contributing factor when the individual's relative efficiency with which new information is acquired, and relatively permanent adjustments made in behavior or thinking, are inconsistent with mission demands. Thus, if training deficiencies are present, the likelihood of errors or of the inability to successfully complete a task increases. What's more, the overall risk to the crew, the vehicle, and the mission increases.

Substantial evidence supports the claim that inadequate training leads to performance errors. Barshi and Loukopoulos (2012) demonstrate that even a task as carefully developed and refined over many years as operating an aircraft can be significantly improved by a systematic analysis, followed by improved procedures and improved training (see also Loukopoulos, Dismukes, & Barshi, 2009a). Unfortunately, such a systematic analysis of training needs rarely occurs during

the preliminary design phase, when modifications are most feasible. Although operational tasks are executed in mockups and simulators by spaceflight crews during preflight training, and feedback regarding interfaces, tasks, and operations is received, often hardware, system and software designs are relatively mature. Therefore, it is often cost prohibitive to modify design based on feedback from training. Hence, designers must consider training when designing hardware and software with crew interfaces, procedures, and operations. Unfortunately, training is often seen as a way to compensate for deficiencies in task and system design, which in turn increases the training load. As a result, task performance often suffers, and with it, the operators and the mission. On the other hand, effective training can indeed compensate for such design deficiencies, and can even go beyond to compensate for our lack of knowledge or for failures of our imagination to anticipate all that might be needed when we send our crew members to go where no one else has gone before. Thus, incorporating training considerations in system design and designing effective training can go a long way towards mitigating mission risks.

References

Barshi, I., & Loukopoulos, L. D. (2012). Training for real world job performance. In A. F. Healy & L. E. Bourne, Jr. (Eds.). *Training cognition: Optimizing efficiency, durability, and generalizability* (pp. 287-306). New York: Psychology Press.

Billings, C. E. (1997). Aviation automation: The search for human-centered approach. Mahwah, NJ: Erlbaum.

Chandler, F. (2007). *Learning from NASA mishaps: What separates success from failure*. Report presented to the Project Management Challenge, February 7, 2007. Available from: http://spacese.spacegrant.org/Failure/Reports/Chandler_Faith_LearningFromFailures.pdf

Dismukes, R. K., Berman, B., & Loukopoulos, L. D. (2007). *The limits of expertise: Rethinking pilot error and the causes of airline accidents*. Burlington, VT: Ashgate.

Department of Defense (2005). Department of Defense human factors analysis and classification system, a mishap investigation and data analysis tool. Washington, D.C.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University. (Original work published 1885)

Ellis, S.R. (2000). Collision in Space. Ergonomics in Design: The Quarterly of Human Factors Applications, 8(1), 4-9.

Endsley, M., & Garland, D. (2000). *Pilot situational awareness training in general aviation*. Proceedings from: 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society, 2000; Santa Monica, CA: Human Factors and Ergonomics Society, 2.357–2.360.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors, 38*, 232-250.

Evans, C. A., Robinson, J. A., Tate-Brown, J., Thumm, T., & Crespo-Richey, J. (2009). *International Space Station science research accomplishments during the assembly years: An analysis of results from 2000-2008*. NASA Technical Paper TP-2009-213146.

Grote, G. (2009). Management of uncertainty: Theory and application in the design of systems and organizations. London; Springer.

Healy, A. F. (2007). Transfer: Specificity and generality. In H. L. Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 271-275). New York: Oxford University Press.

Healy, A. F., & Bourne, L. E., Jr. (Eds.). (2012). *Training cognition: Optimizing efficiency, durability, and generalizability*. New York: Psychology Press.

Healy, A. F., & Bourne, L. E., Jr. (Eds.). (1995). *Learning and memory of knowledge and skills: Durability and specificity*. Thousand Oaks, CA: Sage.

Healy, A. F., & Bourne, L. E., Jr. (1995a). Durability and specificity of knowledge and skills. In A. F. Healy, & L. E. Bourne, Jr. (Eds). *Learning and memory of knowledge and skills: Durability and specificity* (pp. ix-xiv). Thousand Oaks, CA: Sage Publications.

Healy, A. F., Kole, J. A., Buck-Gengler, C. J., & Bourne, L. E., Jr. (2004). Effects of prolonged work on data entry speed and accuracy. *Journal of Experimental Psychology: Applied, 10,* 188-199.

Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2005). The procedural reinstatement principle: Studies on training, retention, and transfer. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 59-71). Washington, DC: American Psychological Association.

Helmreich, R. L., & Merrit, A. C. (1998). Culture at work in aviation and medicine: National, organizational, and professional influences. Burlington, VT: Ashgate.

Horowitz, T. S., Cade, B. E., Wolfe, J. M., & Czeisler, C. A. (2003). Searching night and day: A dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. *Psychological Science, 14,* 549-557.

Kolers, P. A., & Roediger III, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23,* 425-449.

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2009a). *The multitasking myth: Handling complexity in real-world operations*. Burlington, VT: Ashgate.

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2009b). The perils of multitasking. *AeroSafety World*. A Journal of the Flight Safety Foundation, August, 18-23.

MacDougall, R. (1904). Recognition and recall. *The Journal of Philosophy, Psychology and Scientific Methods*, 1, 229-233

Maurino, D. E., Reason, J., Johnson, N., & Lee, R. B. (1995). *Beyond aviation human factors*. Burlington, VT: Ashgate.

McDaniel, M. A., Friedman, A., & Bourne, L. E. (1978). Remembering the levels of information in words. *Memory & Cognition, 6,* 156-164.

McFarland, R. A. (1953). *Human factors in air transportation*. New York: McGraw-Hill.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533. doi: (10.1016/S0022-5371(77)80016-9).

National Aeronautics and Space Administration. Human Spaceflight Gallery [photos on the Internet; cited 2008 Jan 26]. Available from: http://spaceflight.nasa.gov/gallery/index.html.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the power law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

NTSB. http://www.ntsb.gov/ntsb/query.asp

Orasanu, J., & Fischer, U. (1992). Distributed cognition in the cockpit: Linguistic control of shared problem solving, *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 189-194). Hillsdale, NJ: Erlbaum.

Patrashkova-Volzdoska, R. R., McComb, S. A., Green, S. G., & Compton, W. D. (2003). Examining a curvilinear relationship between communication frequency and team performance in cross-functional project teams. *IEEE Transactions on Engineering Management, 50*(3), 262-269.

Proteau, L., Marteniuk, R. G., & Lévesque, L. (1992). A sensorimotor basis for motor learning: Evidence indicating specificity of practice. *Quarterly Journal of Experimental Psychology, 44A,* 557-575.

Rando, C. Baggerman, S.D., & Duvall, L.E. (Eds.) (2005). Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, Aerospace Systems: *Habitability in Space*. Orlando, FL.

Reason, J. (1997). *Managing the risks of organizational accidents*. Burlington, VT: Ashgate.

Rickard, T. C., & Bourne, L. E., Jr. (1995). An identical-elements model of basic arithmetic skills. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 255-281). Thousand Oaks, CA: Sage.

Rickard, T. C., & Bourne, L. E., Jr. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1281-1295. doi: (10.1037/0278-7393.22.5.1281).

Roediger, H. L., III, Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103,* 734-760.

Salas, E., Wilson, K., Priest, H., & Guthrie, J. (2006). Design, delivery, and evaluation of training systems. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (472–512).

Schaafstal, A., Schraagen, J. M., & Berlo, M. van (2000). Cognitive task analysis and innovation of training: The case of structured troubleshooting. *Human Factors*, 42, 75-86.

Schmidt, L., Slack, K., Keeton, K., Barshi, I., Martin, L., Mauro, R., O'Keefe, W., Baldwin, E., Huning, T. (2011, April). *Houston, we have a problem solving model for training*. Presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, Chicago, Illinois.

Shappell, S.A., and Wiegmann, D.A. (2000). "The Human Factors Analysis and Classification System – HFACS", Report DOT/FAA/AM-00/7, February 2000. (http://www.nifc.gov/fireInfo/fireInfo_documents/humanfactors_classAnly.pdf).

Shayler, D.J. (2000). *Disasters and Accidents in Manned Spaceflight*. Chichester, UK: Springer-Praxis Publishing.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Stall, M. A. (2004). *Stress, cognition, and human performance: A literature review and conceptual framework*. NASA/TM-2004-212824, NASA Ames Research Center, Moffett Field, California.

Staal, M. A., Bolton, A. E., Yaroush, R. A., & Bourne, L. E., Jr. (2008). Cognitive performance and resilience to stress. In B. J. Lukey & V. Tepe (Eds.), *Biobehavioral resilience to stress* (pp. 259-300). Boca Raton, FL: CRC Press.

Thorndike, E. L. (1906). *The principles of teaching: Based on psychology*. New York: A. G. Seiler.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373.

Urban, D., Griffin, D., Ruff, G., Cleary, T., Yang, J., Mulholland, G., & Yuan, Z. (2005) *Detection of smoke from microgravity fires*. Proceedings of the International Conference on Environmental Systems, 2005-01-2930

Vogel, R. J., & Thompson, C. P. (1995). The specificity and durability of Rajan's memory. In A. F. Healy, & L. E. Bourne, Jr. (Eds). *Learning and memory of knowledge and skills: Durability and specificity* (pp. 320-342). Thousand Oaks, CA: Sage Publications.

Wickelgren, W. A. (1974). Single-trace fragility theory of memory dynamics. *Memory & Cognition, 2,* 775–780.

Winfred, A., Jr., Winston, B., Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance*, 11, 57-101.

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science, 18,* 133-134.

Wolfe, J. M., Horowitz, T. S., Cade, B. E., & Czeisler, C. A. (2000, November). *Searching when sleepy: Trading accuracy for efficiency*. Paper presented at the 41st annual meeting of the Psychonomic Society, New Orleans, LA.

Team

Vicky Byrne, M.S.	Lockheed Martin, NASA Johnson Space Center
Collin Green, Ph.D.	NASA Ames Research Center
Susan Schuh, M.S.	MEI Technologies, NASA Johnson Space Center
Beth Wenzel, Ph.D.	NASA Ames Research Center

List of Abbreviations

CBT   Computer Based Training

CMO   Crew Medical Officer

DAFT   Dust and Aerosol measurement Feasibility Test

DoD   Department of Defense

FCI   Flight Crew Integration

HFACS   Human Factors Analysis and Classification System

HRP   Human Research Program

ISS   International Space Station

LEO   Low Earth Orbit

NASA   National Aeronautics and Space Administration

NEA   Near Earth Asteroid

NEO   Near Earth Object

NTSB   National Transportation Safety Board

PRD   Program Requirements Document

SA   Situation Awareness

SHFE   Space Human factors Engineering

SHFH   Space Human Factors and Habitability