Relevant Research Assessment Concerning Pilot Response to Unexpected Events

Task 2: Relevant Research Assessment

25 February 2022

Primary Authors; P. A. Hancock[1], Jessica Cruit[1], Janeen Kochan[2], Mira Gruber[1], Yazmin Diaz[1], and Alexandra Kaplan[1]

Subject Matter Experts; Janeen Kochan[2] and Shawn Pruchnicki[3]

University of Central Florida[1], Aviation Research, Training, and Services, Inc.[2], Human Factors Investigation and Education LLC.[3]

This technical report was prepared by the Department of Psychology at the University of Central Florida. This research was completed with funding from the Federal Aviation Administration (FAA) NextGen Human Factors Division (ANG-C1). Thanks to Kathy Abbott (Technical Sponsor) and Bill Kaliardos (Program Manager) for their assistance preparing this report

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Relevant Research Assessment Concerning Pilot Response to Unexpected Events. Task 2: Relevant Research Assessment		5. Report Date February 2022 6. Performing Organization Code: N/A
7. Author(s) P.A. Hancock https://orcid.org/0000-0002-4936-066X Jessica Cruit https://orcid.org/0000-0002-8436-8544 Janeen Kochan https://orcid.org/0000-0003-3783-9048 Mira Gruber https://orcid.org/0000-0001-8645-7906 Yazmin Diaz https://orcid.org/0000-0003-0991-5082 Alexandra Kaplan https://orcid.org/0000-0003-0051-0150 Shawn Pruchnicki https://orcid.org/0000-0002-8758-2255		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Central Florida Department of Psychology 4111 Pictor Ln, Orlando, FL 32816		10. Work Unit No. N/A 11. Contract or Grant No. 692M151940001
12. Sponsoring Agency Name and Address Federal Aviation Administration (FAA) NextGen Human Factors Division 800 Independence Avenue, SW Washington DC, 20591		13. Type of Report and Period Technical report December 2019-September 2020 14. Sponsoring Agency Code ANG-C1
15. Supplementary Notes N/A		

16. Abstract

This report provides a review of the existent information pertinent to the response to novel, unexpected, surprising, and/or unanticipated events, primarily focused on the context of aviation. The primary effort here is to identify ways in which to mitigate the brittleness of accepted traditional forms of response and to foster both adaptive and resilient response capacities throughout the whole of the operational systems. We have examined existing information and have assembled a series of definitions of terms and concepts, primarily revolving around resilient response. We look to knit these terms together and evaluate how the synthetic understanding can be used as a foundational basis for advance. This is a proactive perspective and one that looks to anticipate future threats to aerospace safety to counteract their more adverse influences. The work also provides the foundation for subsequent empirical evaluations of possible challenges by those experiencing unexpected events.

17. Key Words Human factors; human performance; NextGen; resilience; automation; safety; unexpected events; surprise; startle; expertise; training; CRM; adaptability; adaptiveness; brittleness; decision-making; metacognition; robustness; training	18. Distribution Statement No restrictions		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 76	22. Price N/A

Form DOT F 1700.7 (8-72)

Executive Summary

The overarching aim of this project was to provide recommendations for researchers and training instructors to train air carrier pilots on how to manage unexpected aviation events. To this aim, four tasks were completed, and Tasks 2-4 resulted in technical reports. The tasks were as follows:

- *Task 1: Research Plan:* Research team met to create project plan of research.
- *Task 2: Relevant Research Assessment:* Synthesized the extensive literature on pilots' behaviors and responses to unexpected events.
- *Task 3: Pilot Needs Analysis:* Gathered expert input on current pilot performance and feedback to improve responses to unexpected events.
- *Task 4: Training Development Plan:* Recommend training interventions to increase pilot performance during unexpected events.

The present report presents the work of Task 2: Relevant Research Assessment.

This report provides a comprehensive review of the existent information pertinent to the response to novel, unexpected, surprising, and/or unanticipated events, primarily focused on the context of aviation. The need for this work is founded upon the recognition that the numbers of procedural responses to anomalous conditions in flight will eventually prove an insufficient safety strategy in an evolving and complex world of ever-greater uses of automation and autonomous systems. While 'if-then' forms of checklists have provided useful response approaches, the growing number of possible operational states of advanced aircraft mean that attempts at exhaustive state searches are defeated by this computational development. The same structure applies to all who employ ever-advancing autonomous support systems. The primary effort here is to identify ways in which to mitigate the on-coming brittleness of these accepted traditional forms of response and to foster both adaptive and resilient response capacities throughout the whole of the operational systems. The focus here is upon the capacity and capabilities of the flightcrew and additionally we extend this concern to a systems-wide perspective. To accomplish this, we have examined existing information and have assembled a series of definitions of terms and concepts, primarily revolving around resilient response. We look to knit these terms together and evaluate how the synthetic understanding can be used as a foundational basis for advance. We look to distinguish how sudden surprise events, which can elicit startle reactions, might be understood, managed, and trained such that their occurrence in flight, for example, does not prove catastrophic. We acknowledge that these forms of events are liable to become rarer and more idiographic in character across time, most especially as the efficient, protective aspects of automation are themselves perfected. In sum, this is a proactive perspective and one that looks to anticipate future threats to aerospace safety in order to counteract their more adverse influences. The work also provides the foundation for subsequent empirical evaluations of possible challenges by those experiencing unexpected events.

Table of Contents

Figures

Tables

1. Introduction

In order to anticipate the future, one must be cognizant of the past. To this end, our overall project work concerns both resilient and adaptive human-machine responses to coming airspace system challenges and is here founded upon an assessment and synthesis of past relevant work. It covers both topics which are specific to airspace operations, but also embraces, but is not limited to, knowledge, understanding and insight from diverse sources related to performance theory, systems modelling, direct experimentation, and prospective design. This strategy is adopted in order to understand the various ways in which wide and contemporary understanding can be brought to bear on the issues that are identified. While no formal meta-analysis (see e.g., Hancock et al., 2020) of the current area of investigation is possible, we look to provide as precise and as quantitative perspective as is feasible given the present state of the science of resilience engineering.

2. Structure of the Report

The initial sections of this report begin with a brief summary of a previous and important survey concerning the importance of resilience in aviation (Pruchnicki, Key, & Rao, 2019). We then discuss the spectrum of contemporary safety concerns that are faced by those involved in modern aviation. This assessment features both long-term issues that have not yet been effectively addressed as well as new and emerging risks that are presently arising as a result of the introduction of increasing automation and the beginnings of autonomous systems (see Hancock, 2021a). We found our review on an examination of the safety considerations that have been part of aviation's past before exploring why resilience is a critical factor in prospective safety efforts and thus why it is the primary focus of the current report. Since there is a gap in how resilience is assessed and measured during unexpected air carrier events, we explore constructs related to resilience and overall system performance during successful air carrier events. We then discuss uncertainty and unpredictability, and how such concepts are tied to resilience. Following the discussion of unpredictability, we address the facets of expertise which allow pilots to be resilient in the face of unexpected and uncertain events. These expertise factors include domain knowledge as well as judgement abilities. From there, resiliency in action is explored. This focuses on resilience not just as a concept, but as a skill that pilots and other members of the aviation enterprise have used in the past to head off failure. We go on to consider the concept of team resiliency, before explaining the evident difficulties of both measuring and training resiliency. Following these considerations, we evaluate the utility of advances in Crew Resource Management (CRM) and the various ways in which such innovations can mitigate the risks inherent to flight and most especially unexpected circumstances in flight. The report details multiple definitions of the various terms that are central to understanding resilient response. We use these considerations to select which of the definitions are considered to be the most accurate and relevant for this work. We conclude with a purpose-directed discussion of the way resilience, at all systemic levels, can be used to promote greater safety and efficiency in the anticipated and forthcoming Next Generation Air Transportation System (NextGen).

3. A Prior Assessment of Resilience

Our present review is built upon a previous literature review conducted and reported by Pruchnicki, Key, and Rao, (2019) who laid solid foundations upon which the current work is erected. With respect to unexpected events (both those of which that can theoretically be anticipated, and those of which cannot), Pruchnicki and colleagues featured the FAA's (2015) definition of their occurrence as "ones which surprise the pilot by violating expectations." It is circumstances in which standard operational procedures cannot, or do not solve the proximal problem, that we anticipate the need for adaptive and resilient response on behalf of the pilots, the aircraft they are flying, the infrastructure system that supports them and the interactions of all of these levels of response to provide effective resolution strategies.

By definition, where and when off-nominal conditions pertain, pilots must respond to novel challenges that can arise from a plethora of causes. They can be due to security breaches or the origins might lie in some form of mechanical or operational failure brought about by a confusing instrument flight procedure. They might arise from unusual or unprecedented weather conditions, or even issues with the functional capacity of the flightcrew themselves. In some cases, it has been seen that the off-nominal event is related to the mechanistic operation of the aircraft and its operation. The common factor here with respect to required response is that all of these cannot simply be 'trained' for. This therefore mandates a degree of flexible problem-solving and situational response from the pilot (and/or other cooperative members of the airspace system), in order to achieve resolution. These, so-called- 'black-swan' events do not engender a purely algorithmic solution but necessitate innovative solutions.

Flightcrew may be able to anticipate some of these unusual or adverse events by monitoring the aircraft systems to recognize anomalies as they are developing in real time. Such proactive approaches are valuable and are considered extensively in what follows (Rankin, Woltjer, & Field, 2016). However, prospective anticipation, via the use of 'envisioned world' techniques (see Mueller et al., 2019) may well prove insufficient since, as we are aware, many adverse events have proceeded in which pilots are surprised and either never discern the cause of the events they are witnessing, or are not able to distinguish a proper course of action before it is already too late. In fact, these surprise events are one of the most common causes what is deemed later, in hindsight as an operational error (Hassall et al., 2014).

The standard approach to incident prevention and incident response, is the provision of extensive training. This strategy seeks to identify common, and less common, precursors to failure and to inculcate a series of standard responses, anticipated to be effective in resolving them. To be explicit, we advocate, applaud, and adhere to these traditional methods of improving flight safety (some of us having been involved in the delivery of flight safety curricula now for more than forty years). However, not all critical situations can be anticipated (see Hancock & Cruit, 2020). Indeed, according to the FAA, it is simply not possible to prepare pilots for all the novel adverse events which they might encounter (Zuiderwijk, et al., 2016).

4. Maintaining Aviation Safety

Safety is of vital importance in aviation. Traditionally, increasing safety involves the process of mitigating, or ideally eliminating, identified sources of unacceptable risk. While all methods of improving safety have this same end goal, there are a number of different ways through which to achieve increasing safety (Hollnagel, 2014). Two contrasting ideas about safety, which have been labelled Safety-I and Safety-II, advocate for rather different approaches to the question. These are, respectively, founded on different perspectives concerning system behavior and the focus that is featured upon either failure or success. Safety-I is generally directed toward an understanding of failure events and their subsequent prevention. Safety I thus, implicitly at least, adopts the benefit of automated systems. This is because, as each human flaw and failure is sequentially revealed and identified in a system's operation, then an ‘appropriate’ automated counteraction can be envisaged and engaged to reduce human error. This line of thinking leads to increasing automation and shielding of nominally ever-more reliable automated systems. It is the engineering vision of the excision of participation of unreliable human influences. In contrast, Safety-II focuses on the successes of systems’ operations. It emphasizes that systems, such as the airspace one, actually functions exceptionally well because of, not in spite of, the human operator. In this respect, Safety II features the value of human operators, noting that their improvisational and creative nature may well be of exceptional value in responding to unpredicted situations. (Schwarz et al., 2016).

Safety-I often resonates in public opinion and in the political realm, where the narrational need to ‘respond’ to an emergency or failure is responded to as a critical moral imperative. Unfortunately, this can lead to the criminalization of error and is not supportive of an environment where operator error is openly reported and discussed. To date, the implication of Safety I, that we need to forensically understand each failure and then address it, has fostered a line of technological development toward greater automation. It must be acknowledged that this has, across the years, served to make aviation safer. It is thus supportable to assert that the safety and efficiency of flight operations have been increased through the introduction of greater levels of automation. Some of these automated systems have been in use for many years. For many decades we have had degrees of automation involved in flight control. This line of progress continues, in evolving routes and procedures such as Performance Based Navigation (PBN) and other Instrument Flight Procedures (IFPs). IFP’s refer generally to published procedures which specify information such as speed, altitude, and navigation points. These paths can help pilots to avoid obstacles while in some cases simplifying flight deck and ATC procedures. PBN comprised of Area Navigation (RNAV) and Required Navigation Performance (RNP) describes an aircraft’s capability to navigate using specified performance standards. The many types of PBN procedures, covering most phases of flight, have the potential to enhance flight efficiency and safety (FAA, 2020). It is more than probable that these forms of development will continue, especially as they demonstrate proven utility and protection. However, additional automations which can increase safety and efficiency can also serve to increase operator workload (and see Hancock et al., 2020; Hancock, 2013). Instrument Flight Procedures can add to subjective complexity which in turn adds additional cognitive steps for the pilot. It is through this increasing complexity of operations that cognitive load threatens to overwhelm even the most accomplished of flight deck crews (Chandra et al., 2016).

One reason why added cognitive load can threaten or hinder safety efforts is because of the added stress to the pilot (Chandra et al., 2020). This is most especially the case since, for decades now, automation tends to add cognitive workload when task load is already high, and conversely to reduce cognitive workload when task load is ostensibly low. When operational complexity is increased, as occurs when operators are required to observe and manage new technologies, anxiety and stress disrupts normal cognitive operations and diminishes human performance; the example of ‘cognitive tunneling’ being one most particularly relevant reaction here (Hancock & Dirkin, 1983). In order to combat complexity-related decreases in performance, Chandra and colleagues have recommended that pilots develop what they refer to as “*adaptive expertise*” (Chandra et al., 2020). This type of expertise refers to the ability experts have accumulated, whether through training or performance, and apply these learned abilities to new and novel situations for which they have not explicitly prepared. As will be seen, we take some issue with this assertion, especially as it applies to completely new and novel situations. While pilots with this adaptive expertise are asserted by Chandra and colleagues to be better suited to analyzing and responding to new situations in real-time, it is clear that the entire system of operations need to work jointly in this endeavor of fostering a culture of effective expertise.

Across the years of its operation, the FAA has identified many topics as safety issues. In the past decade, issues of prominence have been briefing flight-specific expected threats, surprise, startle, interruptions, monitoring, and cross-checking behaviors (Faerevaag et al., 2018). Briefing flight-specific threats is of use as it is easier for pilots to recognize the signs of a specific and expected issue rather than to focus on the multitude of potential threats to the aircraft that simply might occur. Rivera and colleagues (2014) noted that, while a startle is always a response to a stimulus of high affective intensity, a surprise can come from mismatch of reality with expectations. Here, the latter can include *the unexpected absence* of a stimulus as well the occurrence of an intrusive stimulus. Both terms (i.e., startle and surprise) are frequently used interchangeably by pilots recounting events such as Loss-of-Control Inflight (LOC-I) or a variety of other incidents. In fact, in an analysis of pilot reports where the term "startle" was used, less than half of these individuals were referring to a high-intensity stimulus (Rivera et al., 2014). While these terms then refer to different forms of experience, they can still both result in detrimental influences on aircraft safety and operations.

Interruptions, which serve to stop a pilot mid-task, can prove to be especially problematic. This is particularly the case if the pilot is unable to return to that task in a timely manner. When a task requires a significant degree of attentional focus, an interruption can induce a number of problems, most especially in forgetting or neglecting one or more in a series of required actions (see e.g., Loukopoulos et al., 2001). As well as these attention-focused and intense situations, it is also the case that low-effort tasks tend to be vulnerable to interruptions. Many of the latter forms of task can be carried out without much attentional involvement and here the now added cognitive load of remembering where one left off, adds to the difficulty of resuming the task. Thus interruptions, alongside startle and surprise influences, can act as a source of pilot error (cf., Dismukes et al., 1998). Talone and colleagues (2015) analyzed a number of incident databases in order to determine the relative frequency of incidents involving startle, surprise, and distraction. Their findings showed that distraction was more commonly cited in incident reports than surprise and startle. This latter conclusion was derived from an analysis of 4,773 reports which spanned two decades. Such findings serve to enable us to focus our efforts in proportion to the observed problems. In this case it serves to emphasize the primacy of distraction, an observation that has also been made in other forms of transportation safety (see Hancock et al., 2008).

Martin (2019) has argued that, in almost all aircraft accidents which had some associated, human-related contributory considerations, lack of resilience was an identifiable factor. Martin has noted that such accidents occur when a pilot encounters an adverse event and does not have either sufficient time or cognizance to recover from it. This postulation indicates that both tactical and strategic resilience could have helped the pilot to prevent the outcome which did occur. Martin (2019) further specified one potential cause for this lack of resilience: This was the prevalence of automation in the flight deck (and see also Wiener, 1985). There are two potential reasons why automation may reduce resilience. The first is its lack of transparency that, almost necessarily accompanies increasingly complex and autonomous technology. This can serve to make many states of the aircraft system inaccessible and even invisible to the pilot, representing an overall system weakness. The second reason is the high baseline reliability of such technology, which both increases pilot complacency and decreases their expectation of adversity, turns their monitoring work into an aversive and frustrating vigilance task (Hancock, 2013). Martin (2019) goes on to further note that in previous generations, pilots had higher levels of resiliency, perhaps due to the higher prevalence of threats and potentially due to the more transparent equipment with which they worked. In addition, lack-of, or unsophisticated automation (i.e., automation that was most transparent about what is was doing and it's system state), allowed the pilots to remain more in-the-loop with regard to automation functionality. This might mean that the flexibility inherent in resilience capacity may itself dissolve as a function of increasing automation and increasing peripheralization of the pilot, and the diminishing number of tasks they are required to perform.

One traditional method of maintaining safety can be enacted through expertise. Situations that threaten safety most frequently provide a high degree of stress for the flightcrew. Stress often serves to impair basic cognitive functioning (Hancock & Warm, 1989). It can cause the individual to focus on eliminating the stress rather than focusing on the task at hand; a process often referred to as *attentional narrowing* (Easterbrook, 1959; Hancock & Dirkin, 1983) The higher the stress, often, the greater the degree of change in performance, especially when human operators reach their terminal stage of capacity (Hancock, 2009). If the individual appraises the situation as solvable, they are often able to respond effectively but if the individual believes the situation is, in some way, beyond their capacities, they then experience an excess of anxiety which itself hinders their ability to perform effectively. In general, individuals with greater expertise, when faced with a similar stressful event, tend to manage the situation with more ease than novices. This is why we train in the first place, and as is advocated in the present approaches. As Hancock (1986) put it: "experience with the stress provides protection against that stress, experience with the task serves that same purpose." These and allied observations led Hancock and Warm (1989) to the original postulation that tasks themselves are the proximal form of stress to the active central nervous system and so can be expected to exert the same form of influence as other, external environmental sources of stress (Hancock & Warm, 1989). Well-learned information is often found to be less vulnerable to stress effects than novel information, where even experienced individuals tend to return to novice status. This is one of the relatively hidden challenges of unexpected events, they tend to negate part of the person's experience as a conduit to providing effective resolution. However, making procedures habitual can potentially prevent aviation errors when under stress. (Dismukes et al., 2018).

5. Uncertainty and Resilience

One of the critical concerns that represents a foundation for our overall project is the degree to which the introduction of new and more complex technologies necessarily induces a greater number of unanticipated and unanticipatable events into future operations. For, if this is the case, we cannot use ever increasing degrees of training or even rules to prevent such incidents since the spectrum of possible adverse states may grow commensurately with each technological advance. Safety here is not based then on a search for an exhaustive (and exhausting) repertoire of discrete response patterns, but rather it must feature an approach in which pilots train on improving the types of flexible response skills which can help them respond appropriately in novel events. In this way, the problem looks to recapitulate the empirical exploration concerning the efficacy of massed versus distributed practice in the assimilation of cognitive and motor skills (and see Adams, 1987).

According to Pruchnicki et al., one of the most critical of response skills to these unanticipated events is that of resilience. Resilience is considered as a capacity that serves to improve an individual's or team's response abilities, so that when they are faced with a novel situation, they are able to react in an effective manner (and see Dekker & Lundstrom, 2006). Pruchnicki et al. (2019) noted that individual, human resilience is a set of abilities which can be trained for. So, while the precise and deterministic response chain in reaction to adverse events cannot be taught, due to the proliferation and variability of problems, the skills to cope with such events can be learned. Again, there is a parallel in the motor control literature in which this difference is emphasized between the enactment of specific motor programming versus the activation of more general motor schema (Schmidt, 1975). This personal process is one element of resilience engineering which helps to teach the overarching skills that can be applied in many diverse situations (Dekker & Pruchnicki, 2013).

The four Cornerstones of Resilience according to Hollnagel (2009), are (a) anticipating, (b) monitoring, (c) responding, and (d) learning. Anticipating refers to making frequent assessments of the situation, so that any potential threats, which are not yet manifest, can be prepared for or at least anticipated. Monitoring refers to keeping a close watch on equipment so that malfunctions or other issues can be detected promptly (and see Hancock, 2013). An important component of this is knowing what to monitor. Responding is the action that takes place, in reaction to the event. Finally, learning refers to the fact that, after an event has occurred, it is no longer novel. Other similar events in the future can then be adequately prepared for once the situation and the outcome are known (Hollnagel, 2009). These cornerstones represent skills which can be honed through resilience engineering and associated resilience training.

6. Why Resilience?

One of the primary issues here is why resilience is such an important safety feature, as compared to the many other allied constructs which might also improve the safety of aviation. After all, technological advancements, personnel management strategies and programs, improved automation, and enhanced weather forecasting can all help to mitigate prospective operational problems. Resilience is, however, most especially useful in situations where unexpected events occur and associated operational errors may not be expected or threaten to be unrecoverable. Under these conditions, resilience may appear to be useful only in specific circumstances. However, almost all extant safety enhancements only prevent problems that can already be known (and see Hancock & Cruit, 2020). Pilot resilience is concerned with increasing the probability of a more favorable outcome during the noted unexpected incidents. As systems become more autonomous as well as more complex, and iterative design helps to prevent each of the known problems, some incidents will still arise from problems that are, nominally at least, unknown (and see Hancock, 2021a). For example, better weather monitoring may decrease the problems associated with this inherent operational uncertainty. Increased automation can decrease issues of fatigue and human error. Resilience is, however, the principal tactic which can respond to all these problems, as well as problems that have never been encountered before. Therefore, resilience training is a *proactive* way to prevent incidents, as opposed to the more common *reactive* methods which prevent incidents from being repeated but not from occurring in the first place.

7. Traditional Expertise Literature

Experts are distinguished from novices by their knowledge and/or skill in a specific field of endeavor (Cruit, 2016; Ericsson & Lehmann, 1996). Expertise is often conceived as being domain specific, in that expertise in one field does not necessarily translate to expertise in another. This degree of potential transfer, of course, is contingent upon the degree to which the specific domains of concern overlap. This overlap can be either in terms of the cognitive or the psycho-motor demands of the work at hand. Expertise is not only an aspect of an individual but proves most frequently, more specific to a certain human-situation interaction (Cruit, 2016). On the topic of assessing expertise, Cruit (2016) has noted that it is important to use measurable objectives as requirements for the demonstration of expertise when using standardized assessments to measure the level of an individual’s capacity to respond. To a certain degree, expertise can be reproduced in an artificial setting, such as a laboratory or a simulation environment, but standardized methods must be used to measure performance in each of these cases.

To measure success in the laboratory by one metric, and success in the real world by another, is neither a valid nor an advisable strategy.

Cruit (2016) goes on to posit that there are two common characteristics that domain experts generally possess. First, experts are faster than non-experts at perceiving, recognizing, and retrieving informational cues (and see Ericsson & Smith, 1991; Ericsson & Lehmann, 1996). Second, an expert can link information together to come to a different and even novel understanding of relevant, yet unlearned material (Chi & Glaser, 1988; Ross & Spalding, 1991). Expertise accrues gradually, over time. The more system-wide expertise that is present, then ostensibly the safer any task or process becomes. This is because quicker cognitive processes allow for recued response latency. In operational realms in which time is nearly always an important, and sometimes a crucial factor, improved response latency often represents the barrier between failure and success. Experts retrieve their knowledge from long-term memory whereas non-experts tend to use short-term memory to try to resolve proximal challenges. Using short-term memory to assess and interpret a situation takes more time, increases response time while also increasing the associated level of cognitive workload. (Hancock, et al., 2020). Short-term memory also limits the amount of information being processed at once. Since experts also access long-term memory, they are able to process more information at a faster rate. Experts are also able to categorize information and create patterns to distinguish between usual and unusual situations when given new information. Experts create mental simulations/situations and use them to predict future situations (Cruit, 2016). In this sense, and most especially in relation to being able to 'see' and adopt new paths of response, experts prove, most often, to be more resilient than their less skilled peers.

This classical view of expertise suggests that aggregation of experience (e.g., ten years of full-time work in a domain) is the single most important factor in the acquisition of expertise (Chase & Simon, 1973). In most cases, no distinction is made concerning the type of experience the expert has had as long as it is acquired roughly in the domain of interest. On the other hand, studies by Ericsson and Smith (1991), Libby and Frederick (1989), and Gustafson (1963) revealed that people with many years of experience in a domain performed only slightly better than those just coming out of training. They conclude that the greatest amount of improvement occurs in training, not as a result of years of experience. Therefore, a more granular view of pilot expertise, and what factors contribute to it, will be considered next.

7.1 Domain-Specific Expertise

Although expertise is considered predominantly domain specific, general attributes of cognition contribute to the level of performance by experts. This has proved most especially true in aviation. Kochan (2005) investigated the constructs of domain expertise and judgment expertise in studying pilots' reactions to unexpected events. Domain expertise was found to be indexed by one's declarative, procedural, and structural knowledge. These respective relationships are depicted in Figure 1. *Declarative Knowledge* is the number and availability of facts to which an individual (or more recently, a system) may have access. *Procedural Knowledge* is the number and availability of procedures which may be used in domain specific tasks. *Structural Knowledge* is adeptness at how to accomplish a task. Classic research has affirmed that people who have developed expertise in an area show differences from novices in how their knowledge is stored, structured, and subsequently retrieved (Chi et al., 1988).

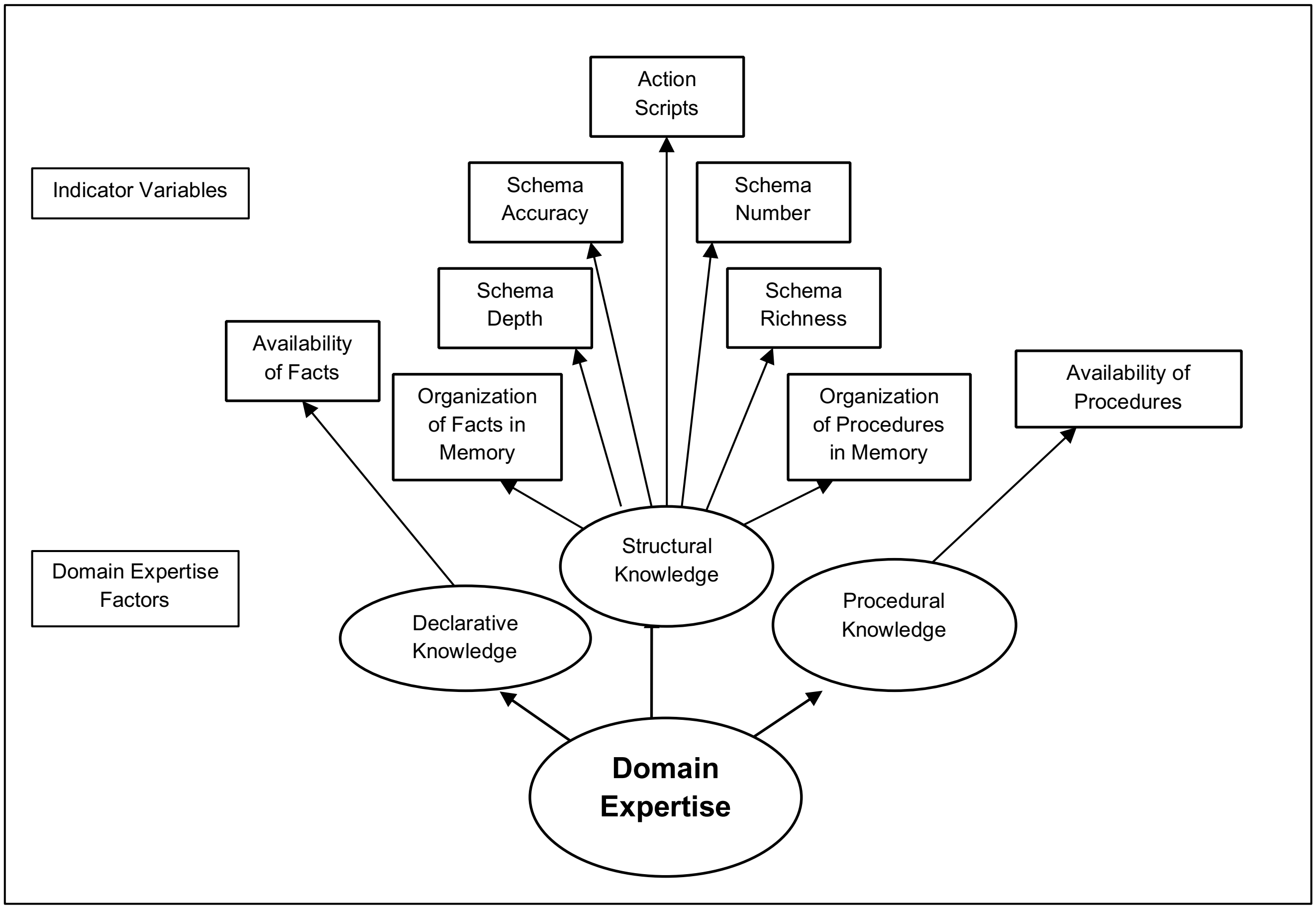

Figure 1. Domain specific factors involved in domain expertise (Kochan, 2005).

We can locate these distinctions between novices and experts using a series of supportable assertions and these are given below. The most consistent and observable differences in the organization and structure of knowledge found between experts and novices across domains include (but need not necessarily be limited to):

1. Experts notice features and meaningful patterns of information that are not noticed by novices.

2. Experts use forward reasoning and incorporate more intuition in problem solving than do novices.

3. Experts are better at ignoring irrelevant cues in the environment than are novices.

4. Experts are better able to incorporate relevant aspects of the situation with their existing knowledge than are novices.

5. Experts' knowledge is extensive and is organized in a way that reflects a deep understanding of the subject matter, a concept which was modeled mathematically and called *compiled knowledge* by Anderson (1983).

6. Experts have varying levels of flexibility in their approach to new situations.

7. 7. Experts are not necessarily adept at teaching, even in their domain of expertise.

An expert's organization and breadth of knowledge (item 5), in one respect, facilitates efficient and effective problem solving. They have more conceptual chunks available to both long and short-term memory, more cues defining each chunk, more interrelationships among the cues and chunks, and more pathways available to access all, or parts of the chunks as needed, to solve a problem or make a decision (Klein, 1993). In this sense, with differing forms of deterioration, expertise can be lost. Such losses can be due, for example, to aging and failures of memory. However, the erosion of expertise might also result from the ever-greater role of automation such that the reliable action of such automated aids can help in the process of forgetting; as has been explored in general in psychological research across the past century (see e.g., Bartlett, 1932). What is not consistent among experts is their ability to be flexible and use cues appropriately in unexpected or surprising situations. This is where the concept of judgment expertise becomes of concern and so it is to this that we now turn.

7.2 Judgment Expertise

The concept of separate constructs of domain expertise and judgment expertise is integral to the identification of the specific underlying skills necessary for responding to unexpected events. It is the tight integration of type, amount, frequency, and diversity of training, and experience, intertwined with other cognitive skills and abilities that need to be resolved to determine what and how to improve pilot performance. To this end, we parsed out the relative contribution of domain expertise to one's performance on the task of dealing with an unexpected event in order to identify the role of judgment. Conversely, other research has attempted to separate the judgment skills away from the domain expertise by minimizing the necessity of using judgment in the completion of tasks. Studies conducted in assessment centers (Kleinmann, 1993; Smith-Jentsch, 1996) indicated that revealing the nature of the domain task (making the dimensions of the task transparent to the applicant) improved the construct validity of the task.

Certainly, specific domain expertise plays one role in successful responses to surprise. Therefore, we believe that merely increasing one's level of domain expertise or domain knowledge is not the most effective way to improve responses to unexpected events. By reviewing general judgment and decision-making literature, the following model (Figure 2) of judgment expertise was developed.

The components of *judgment* expertise are presented in Figure 2, and we here explain each of the constructs noted.

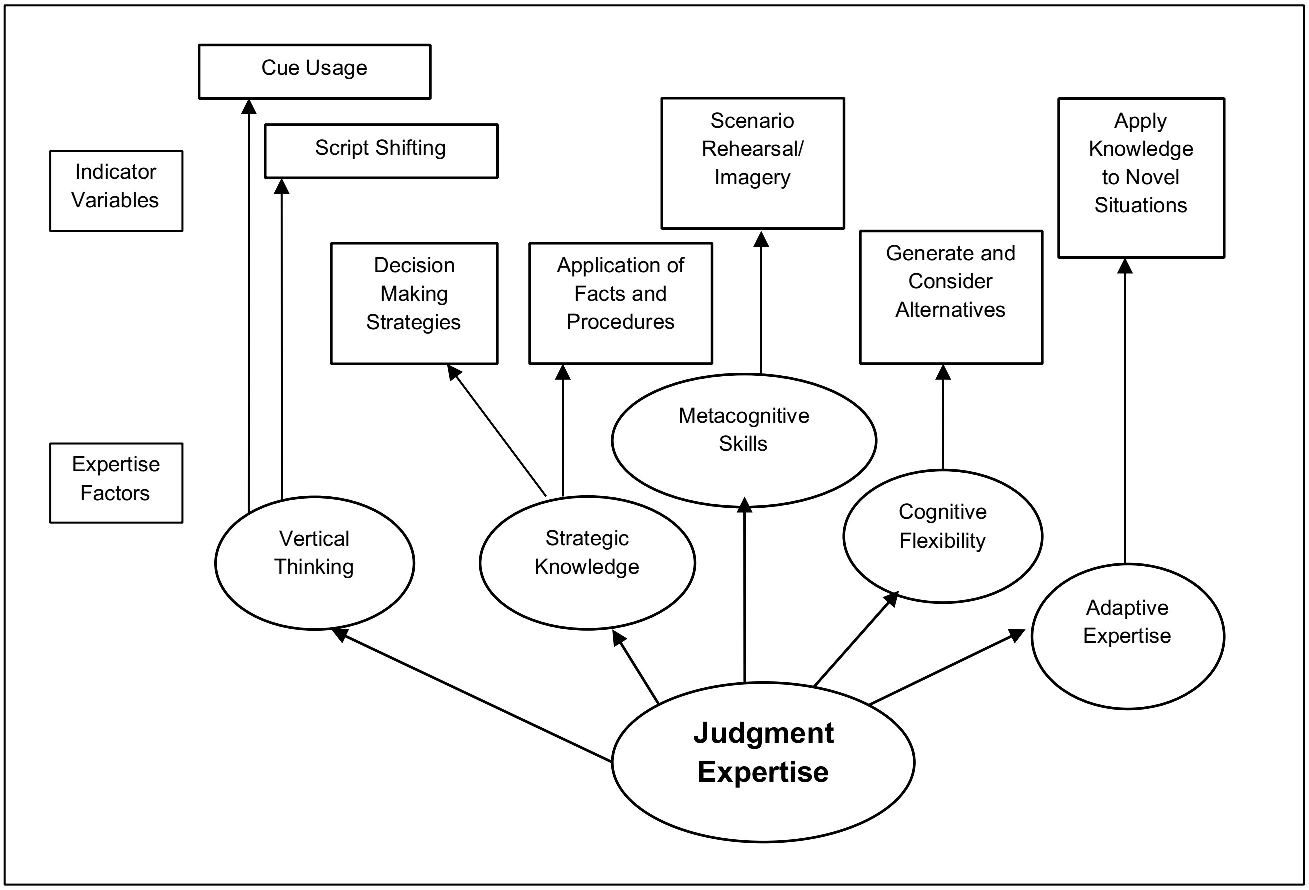

Figure 2. Domain independent factors involved in judgment expertise (Kochan, 2005).

Strategic Knowledge is the compilation and application of facts and procedures and the decision-making strategies used in the process. *Decision Making Strategies* provide an exhaustive review of information processing and decision-making theory is not warranted here since a number of contemporary sources have discussed and examined these extensively (see Kahneman, 2011). However, we briefly consider one pervasive idea found in judgment and decision-making literature; that there are two pathways available in decision making; the analytic and the intuitive (Kahneman & Klein, 2009). Often, the delineation between these modes of information processing and response is based on the amount of attention and cognitive energy required to make a judgment and respond to a given stimulus or situation. This short excursion is necessary to the understanding of the process of surprise and how it relates to resilience because cue (information) processing will depend on task condition variables, such as time available, uncertainty in the information, level of expertise, and one's cognitive flexibility.

In this light, Kenneth Hammond's (2000) *Cognitive Continuum Theory* posits that decisions are made in response to a disruption in constancy, and that optimally, one would move along a continuum from intuitive to analytic processes as the need arose due to the changing nature of the task. The properties of intuitive decision making are low cognitive control, fast processing, low attentional energy consumption, and low conscious awareness. Analysis incorporates high cognitive control, slower processing, high conscious awareness, is task specific and when errors occur, they are few, but large. Hammond also states that it is often an unexpected event, which demands immediate decisions and responses that drive decision making toward the intuition end of the decision-making spectrum.

The view that humans process internal and external sensory cues from an *n* dimensional, hierarchical pattern array; employing a synchronous method is the basis for what we have called *vertical thinking*. Each level in the hierarchy aspires to provide more (higher levels) or fewer cues (lower levels) in a representative pattern. Each cue *pattern* contains varying percentages of relevant and/or irrelevant cues which provide input to the processes used to choose an appropriate action from candidate response sets. Each *cue* varies in its interrelationships (e.g., excites or inhibits) and in its strength of interconnectedness with other cues. Thus, the ability to quickly and accurately assess a surprising situation, process the relevant cues, select an opposite response, and then initiate (or decline) action is dependent on one’s ability to target the most useful level of the overall array of cues presented.

Cognitive Flexibility addresses the pilot’s capacity to restructure the current action script quickly and accurately in order to adaptively respond to dynamic situations and thus, assumedly, be better able to react to unexpected events (Spiro et al., 1988). According to Hammond (2000), cognitive competence in this context, can be considered of a two-fold nature. These are, subject matter or domain competence and, separately, judgment and decision-making competence. Domain competence is a function of learning, memory, and deduction, while decision making is the execution and application of the acquired domain knowledge. It is this idea of decision-making competence, when coupled with domain expertise that gives rise to *adaptive expertise*. *Metacognition* is the ability to monitor one's current level of understanding and decide when it is and when it is not adequate. In other words, it is the awareness of one’s knowledge and is a skill which can be used to control and manipulate cognitive processes. In order to develop adaptive expertise, pilots need to understand how they think, and how what they currently know can be helpful, but at times can be detrimental.

7.3 The Expert Pilot

When considering the human attributes needed to be resilient in the changing environment of aviation, the following enumeration of facets that contribute to pilot expertise can be considered. Kochan et al. (1997) found that expertise in pilots can be defined in terms of the following ten characteristics (and Figure 3) demonstrates how each of these capacities fit into a model of an expert pilot.

1. Self-confidence in his or her skills as a pilot.
2. Motivation to learn all there is to know about the flight domain and practices their skills constantly.
3. Ability to focus on the necessary task and change that focus at the slightest hint that a change is needed.
4. Situation awareness (flight environment, location of other aircraft, terrain, navigation, communications, weather, etc.).
5. Cognizant of the machine including noise, vibration, and engine indications.
6. Vigilant for the unusual, abnormal, or emergency, and mentally makes contingency plans.
7. Mental capacity for problem diagnosis, risk assessment, and problem resolution.
8. Communication skills and applies those skills to each audience and situation.

9. Knowledge of his or her own limitations and motivation to keep a safe margin above those limits.

10. Ego-strength to enforce his or her own limitations in every situation.

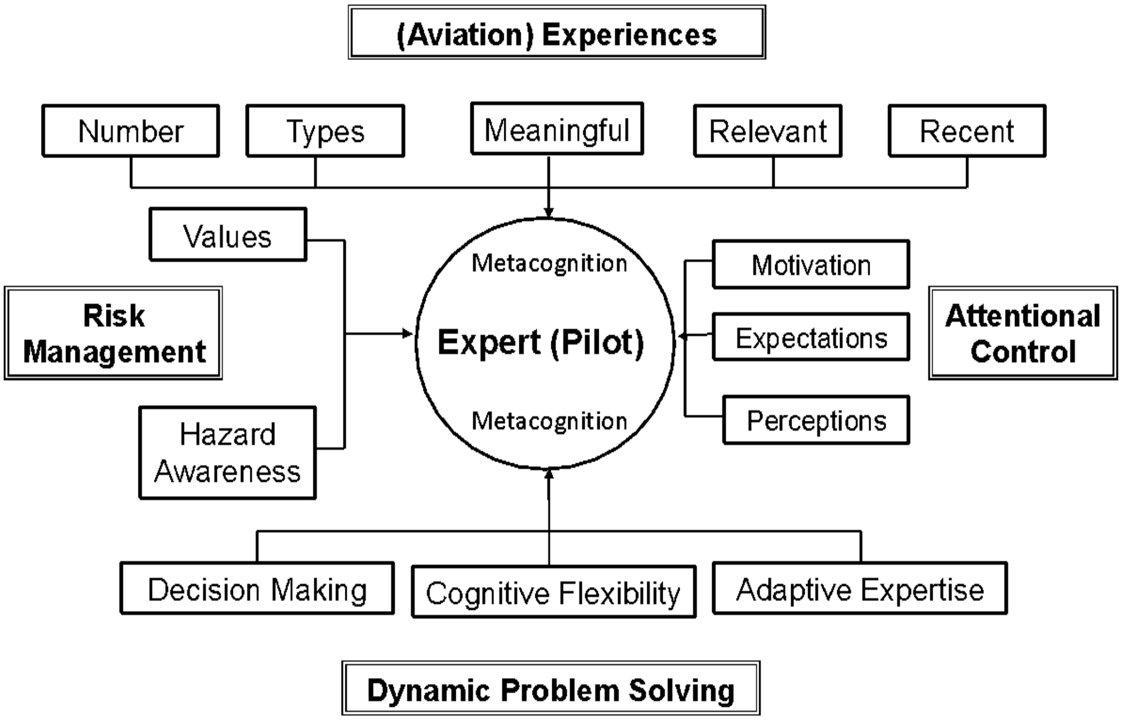

Figure 3. Attributes of pilot expertise adapted from Jensen (1995).

8. Resiliency in Real-World Events

Recall that Hancock and Cruit (2020) define resilience as ““the capacity of a system to exhibit a new state of operational stability when adaptation to recover the prior base state has now failed” (and see Appendix A for more definitions).” In contrast, Faerevaag and colleagues (2018 define resilience as the “capacity of a system to absorb disturbances and adapt to a change in order to retain the same function, structure, and identity.” This latter identification by Hancock and Cruit (2020) is, rather naturally, ddirectly aligned with the conceptual identification advanced by Hoffman and Hancock (2017). Faerevaag’s (2020) definition posits that a capacity of resilience is retaining the same structure as before the disturbance, that is a return to the fundamental configuration that was in place before any disturbance was experienced. Hancock and Cruit’s (2020) definition identifies the additional capacity that resilience requires the ability to change to a form and/or function that better fits the needs brought about by the presence of said threat or disturbance. While these differences may seem small and even obscure, the latter identification reveals that true resilience is the ability to operate in more than one state. Rather than solely attempting to return to one, putatively ideal state, which Hoffman and Hancock (2017) identify as the exercise of adaptability, in the Hancock and Cruit (2020) approach, should that effort at adaptability prove insufficient, then functioning in a new manner is preferable to continual and ultimately futile attempts to reverse the changes wrought by any outside or inside threat. The European Aviation Safety Association (EASA) (2017) has their own definition which does not contradict either of the noted classifications. The latter group sees resilience as “an ongoing and adaptable process including situation assessment, self-review, decision and action.” This is a descriptive specification of general response processes and bears a rather obvious similarity to approaches describing situations awareness (SA) (Endsley, 1995; Smith & Hancock, 1995). In line with the definition of SA, EASA also features the prospective dimension of anticipatory response; although the role of prospection in relation to unexpected and unanticipated events is a complex one, as we discuss later.

The Civil Aviation Authority (CAA) in the UK found that between 2010 and 2011, flightcrew mistakes were the primary cause of over half of aviation accidents analyzed. Of these accidents, non-technical skills accounted for 32% of these actions, while errors in managing the aircraft accounted for only a total of 14%. Thus, they concluded that crew skills, both technical and non-technical, require the type of improvement that can come from enhanced training (UK CAA, 2013). EASA noted in their 2014-2017 European Aviation Safety Plan (EASP) that aircraft crew could be trained towards the competencies needed to keep them safe via both evidence-based, and competency-based training methods. Their safety plan for 2016-2020, revised and then called the European Plan for Aviation Safety (EPAS) also made note of the new issues that would arise in aviation, and that flightcrew members would need to be trained to prepare for these modern challenges (EASA, 2016). It is also worth noting that such resilient capacities are not confined to aviation but are being advocated for many industries dealing with safety-critical systems.

The new challenges that have been identified by EASA among others, include the lack of experience that many pilots now have in manual flight control. With automation conducting much of the work previously completed by pilots, flightcrew manual skills can begin to degrade and dissolve. Worse, some of these skills can become viewed as vestigial and therefore practiced little if at all. Without practice handling the aircraft, pilots' abilities to respond when needed, such as in the case of automation failure or unexpected threats, is potentially compromised to a problematic degree (and see Frazzini, 2001). Therefore, efforts that seek to modernize training must focus on use of, and teaming with, the automation. It must also emphasize the retention of rarely utilized manual flight skills, all the while acknowledging the change in control form inveighed by automation's respective advances (EASA, 2016). Human factors research has focused and reported on how automation can assist pilots, without replacing them, in detecting and reacting to unexpected events that threaten safety (and see Landman et al., 2017). EASA (2016) also pointed out that, while aviation is generally safe, it is not without risk. In fact, the high level of safety means that those rare cases of adverse events become sequentially less predictable than they ever have been (and see Hancock, 2020). The focus on safety in aviation has led to many improvements, through iterative design, to any aspect of flying that has involved what has been perceived as an unneeded degree of risk. Therefore, wherever possible, any 'known' causes of errors have been mitigated. This leaves only the *unknown* risks, which are, by nature, impossible to anticipate (EASA, 2016; and see Hancock & Cruit, 2021a).

Those incidents that do occur are now, most often, not due to simply to one singular source. If that were the case, that singular source would have been determined, and improved, as aviation technology has progressed. Many modern adverse events involve human and/or automation factors. Since these types of events usually involve two or more involved entities, they are hard to predict and prevent. Human pilots can only adapt to respond to and be resilient with respect to these concatenated failures. Here, the EASA (2016) report observed that human adaptability is the capacity that still makes people better pilots than automation at this present time.

The EPAS, derived from the EASA (2016) report, separated the challenges facing aviation into three overall categories. The first were the systemic issues, which are problems to do with the entire aviation system such as issues relating to individual humans, or to cultures (whether those problems are company-based or industry-wide). The next category involves operational concerns, such as weather, runway problems, and other operational and logistical issues which have often been the cause of incidents across the history of aviation. The third category is represented by new and emerging issues. These are threats that did not exist in prior generations, such as problems and challenges associated with Unmanned Aircraft Systems (UAS) and cybersecurity. We are once again reminded of Martin's (2019) work and more specifically, the idea to consider the multi-layers of resilience (i.e., individual resilience, team resilience, organizational resilience, and systemic resilience.

The overarching European Commission (2018) has itself, identified five emerging challenges which threaten to make aviation potentially more dangerous. One of these challenges is automation. The challenge is the increase in automation in both the flight deck, the air-traffic control facility as well as similar advances in Airline Operation Centers (AOCs). While automation is generally conceived of as contributing to a safer environment when it is working correctly, there are times at which it can fail, and then even perniciously so. When it does fail, the human operator is left to control the situation, after having been removed, almost totally, from the often-opaque automated process. The human must then work in order to determine what has happened, up to the point of the incident, as well as dealing with the incident itself and its differing impacts. This can often prove to represent an excessive level of task demand for one individual, especially as there is often little to no warning until it is close to being too late.

The European Commission (2018), after reviewing the research related to aviation incidents, made several recommendations for aviation policy. Many of the recommendations were to do with what they called *no competition on safety*. The idea was to improve safety by collaboration between agencies and companies that would normally be, if not competitors, then almost completely separate entities. They suggested sharing any relevant safety items of information such as new inventions and improvements. Additionally, the Commission suggested sharing research and data that could lead to any of the previously mentioned safety items. In this way they encouraged dissolving the differences between safety and security to focus on one joint venture of making aviation less risky. Another suggestion was the greater utilization of human-factors and risk-informed research. Additionally, they suggested laws and regulations relating to new technology such as UASs and Artificial Intelligence (AI) before such technologies are fully integrated into the aviation industry; this rather than engaging in damage control and limitation after any of their adverse effects had become manifest (European Commission, 2018).

9. Team Resiliency

To the present juncture, we have focused on the character of resilience as it pertains to the individual human member of the aviation systems. To a lesser extent, we have featured resilience as it relates to the technological support systems that foster and permit safe and efficient aviation operations. However, resilience can be expressed at levels of analysis other than the individual and it is to this dimension of team resilience that we now turn. As a general prospect, team resilience is more complicated because it involves multiple individual and/or associated entities, some of whom possess more resilience than others (Martin, 2020) and may be working independently or specifically together. While there is little research on the topic of team resilience, some information is presently available. In these cases, it has been suggested that a strong leader proves to be pivotal (Martin, 2020). One potential way to achieve team resiliency is through joint training. Joint training is something relatively new and something which, according to Faerevaag and colleagues (2018), has not been implemented in many of the present US air carrier training programs. This joint training helps build trust amongst crew members and facilitates communication (which is a commonly observed and advocated CRM skill) and coordination between them. Some instructors have even stated that joint training has prevented misunderstandings between pilots and crewmembers (Faerevaag et al., 2018).

Just as joint training can improve teamwork, so can shared prior experiences of emergency situations. Martin (2019) has noted several instances where near disasters were averted by team resilience, rather than any one individual's specific response abilities. One such example is the United Airlines Flight 232 on July 19, 1989, in Sioux City, Iowa (National Transportation Safety Board, 1989) where a complete loss of hydraulic pressure and an engine failure nearly led to an immediate crash. However, the flightcrew, which included a captain who was merely a passenger on the aircraft, were able to delay the crash by manipulating the thrust levers until the aircraft was closer to the ground. While many people lost their lives during the crash, many were saved. Without the recorded expression of team resilience, it is more than likely that all onboard would have perished. This is especially the case had not the initial problem been controlled.

10. Training Resiliency

Hancock and Cruit (2020) have stated that it is especially important to consider when training interventions are adopted. They also noted that there are three possible sources of resilience. The human pilot may express resilience, as can the automated system with which they are interacting. However, this interaction, between human and automation can be a third source of resilience. While in previous iterations, the focus of resiliency training has always been specifically on humans, the new era of human-machine interaction necessitates a focus on machine resilience also. Resilience within a system can be implemented at the micro, meso, and macro levels. The micro level may refer to resilient individuals, whether those are pilots, other flightcrew, or an individual automated machine. The meso level may refer to a resilient group, whether that is the interaction between pilot and machine, or the community of a certain air carrier. Finally, the macro level, or global level, would refer to a larger group such as all air carriers. Hancock and Cruit go on to suggest that training should take both a top-down, and bottom-up approach as focusing only on the micro level may well later serve to cause problems at the macro level (Hancock & Cruit, 2020). While these suggestions do not answer the question of how to measure and train for resilience, they point out some important concerns for the ways that training is currently being conducted (cf., Robertson et al., 2015).

While it is possible that expertise may well be a helpful foundation for resiliency, these two constructs are not necessarily equivalent. While novices have a lack of experience which may complicate their reactions in novel events, prior (general) experience is not always necessarily an indication that a pilot will be resilient, or even that they will have an advantage over more novice pilots. Cruit et al. (2019) found that even pilots with greater numbers of flight hours can still be significantly challenged during flight. More specifically, the authors found that pilots who held lower flight hours were still able to leverage prior experiences to successfully manage an unexpected event, while some pilots with higher flight hours were not able to leverage past experiences. This suggests that when considering training, it is important to not only focus on flight hours, but also training the pilots on how to associate prior experiences with novel situations. It is thus important to establish correlation matrices indicating the strength of the relationship between these expressions of response capacity. Some, such as flight hours, are obviously objective values and relatively easy to establish. Others, such as resilience capacity remain, at this time, more nebulous and difficult to capture. However, responding to unexpected events and the way it relates to prior personnel experience represents an important step that our project is examining. Other potential variables which might serve as proxies for resilience are pilot self-efficacy, and cognitive flexibility, as we explore below.

EASA (2017b) claims that CRM uses the resources of humans, equipment, and procedures to increase safety and efficiency and that these skills can be trained. The skills that can be enhanced via education include cognitive, technical, and interpersonal ones. Cognitive skills here refer to decision-making and mental abilities of the individual. Interpersonal skills refer to their ability to communicate effectively with each other as well as other personnel across the whole aviation system. Technical skills involve the ability to use automation, complex systems, and other machinery. All of these skills can overlap. Cognitive skills may aid someone's use of technology or inform their interpersonal relations. High skills in the interpersonal area may improve someone's ability to interact with systems or other team members. In general, all three are important (EASA, 2017b). The three skillsets of interpersonal, cognitive, and technical abilities can help to prevent error. Another way to increase safety is to promote an environment wherein communication between pilots and other crew is open and unrestricted. Risk increases when errors remain undetected and/or unaddressed. Automation has improved aviation safety; however, automation must adhere to aircraft design and philosophies of operations. The use and acknowledgement of automation has to develop along with the evolution of aircraft design. As well as workload increase due to automation, unexpected occurrences such as surprise, and startle events, can decrease safety by diverting operator attention from the task at hand.

We discuss the aspects of startle and surprise in more detail below. However, startle frequently evokes the physiological fight or flight response. When recovering from a startle that occurs during operations, it is imperative that pilots continue in actively 'flying' the aircraft and only take immediate actions when they become necessary to do so. As far as is feasible, pilots should refrain from impulsive actions as they have been found to be counterproductive and sometimes dangerous. It is recommended that pilots overlearn recovery responses, to permit implementation without excessive conscious effort, and know the event cues to support faster recognition and response (Martin, 2019). Taking the time to analyze the situation and identify the problem as a crew allows for sound decisions to be made. As startle and surprise can tax pilot resilience and CRM skills, implementation in training may prove useful, if not essential (Martin, 2019). Startle is, of course, often only the first event in a sequence of occurrences, again more of which is discussed below.

EASA has also noted that Threat and Error Management (TEM) involves accepting the fact that problems and errors will occur. It is not possible, no matter how well-trained, for a crew to prevent every possible threat. These threats will occur and to decrease negative outcomes, they must be managed appropriately. Resilience holds the promise to help increase survivability of many actual and incipient incidents. EASA (2017b) noted that resilience is a system's ability to adapt in such a way that it can maintain those processes that are necessary for operation, even during unexpected states (cf., Hoffman & Hancock, 2017). Resilience is often considered a natural skill that some people innately possess but can also be interpreted as the degree of skill level. However, it is the common belief of the research community that training can actually aid in the development of resilience. There are several suggestions put forth by EASA on how to build a good training regimen (EASA, 2017b). While some have argued for the removal of 'resilience development' from training because it is related to personality and personalities cannot be changed during CRM training, resilience development is specific and training topics should be broad to allow operators to mold the programs. However, as of 2020, EASA does maintain a focus on resilience training. As startle and surprise responses can be difficult to elicit in classrooms, the use of Flight Simulation Training Devices (FSTD) are supported to mimic in-air situations (EASA, 2020). The latter avenue has also been explored in relation to ground transportation and the associated startle effects there (see Hancock & de Ridder, 2003).

The idea of training for resilience is somewhat controversial, as it is always problematic trying to train for the unexpected. Indeed, it is this precise process that has been thought to underlie the development of play in young animals who are, effectively, engaging in forms of real-world simulation in order to deal with surprise when it actually does arise (Spinka et al., 2001). Training difficulties here are also partially due to the fact that most threats that require resilience, rather than learned responses, are unknown. In general, what is unknown cannot be trained for and, if it were trained for, the response would not therefore be “resilient” per se but would entail falling back upon extant training. As we have noted, it is difficult to truly startle people in a classroom, or even in a simulator, setting. For this reason, resilience engineering is quite difficult to teach, and any training plan designed to increase flightcrew resilience must be carefully thought out. Luxair’s course on resilience engineering has been made public in order to help improve resilience training programs for all flightcrew (Luxair Human Factors Training Team, 2013). In this they noted that a system able to still perform its required functions while under external, unexpected pressures, is a resilient one. They defined a system as multiple entities, whether the entities are one person and one computer, two people, a crew of people, or a whole company. These systems can be resilient at the individual or team level. They can also possess weak points, again, at the individual or team level, where an entity fails to adapt and becomes stuck in one particular mode of operation, choosing to make incorrect decisions purely because they have worked in the past.

Luxair (2013) also noted that over time, it has become more difficult to determine whether safety training programs actually prove to be effective. This is due to the fact that fewer and fewer incidents occur each year. As safety improves, there are almost no accidents to investigate, and fewer “near misses” to examine as potential successes of safety training programs. Simply put, there is a floor effect when incidents prove to be so rare. As the prevalence of aviation incidents is so small any decrease that accrues from a new safety program, can only be limited in its effect. In many ways this is a function of success. As relative and absolute numbers of incidents decrease, the traditional target measures become less and less effective. This resulted, traditionally, with a greater emphasis on non-disastrous incidents. It was the equivalent of moving one level down Heinerich’s triangle (Heinrich, 1931) or, more polemically, pushing back one layer of the well-known “Swiss Cheese” slice model (Reason, 1990). As we envisage our future, perhaps the target criterion will itself have to migrate further up the causal chain, such that we monitor all forms of deviation in a manner somewhat akin to Demming’s approach to statistical quality control (Deming, 1953). That such an effort can now be accomplished, dynamically and on-line as the vehicle is in the air, holds great promise for even further safety gains.

Every aspect of aviation has an associated “risk profile.” Pilots are aware of these profiles and can prepare for them. For instance, an airport with a difficult approach is an issue that is typically well-known to pilots. They do not begin their landing process without being aware of the difficulties specific to that approach. These difficulties require preparation. However, the types of difficulties that requires resilience are ones that are not known and therefore not prepared for (and see Hancock & Cruit, 2020). Luxair (2013) identifies three conditions under which aircraft personnel can respond incorrectly to an unexpected threat.

1. *Decompensation:* When enough goes wrong, the parts of the system that have been ‘compensating’ or making up for these issues, can get overwhelmed. When this happens, they can no longer compensate and an additional problem, however small, could prove disastrous. The remedy for decompensation is teamwork or allowing *all* aspects of the system to take on some of the burden of compensation so only one part does not get overwhelmed.

2. *Working at cross-purposes*: This refers to parts of the system trying to achieve different goals. For instance, one part of the system may be attempting to keep everything on schedule, while another may be trying to save time by flying faster. It is important to focus on overall success, not individual success. Communication, both within a small system and a larger one, is the remedy to this problem.

3. *Outdated behaviors*: Relying on outdated behaviors is a problem when personnel have a specific view of what could be happening at any given time. They might assume that something is not a threat because it has not been a threat in the past. However, if signs are pointing towards a situation being something other-than-expected, it is important to understand what is happening and to learn from the signs, or to regain situation awareness—for the *real* situation, not the one initially assumed.

The three issues are decompensation, working at cross purposes, and outdated behaviors. The three remedies which are recommended are teamwork, communication, and situation awareness. These remedies point the way towards resilience and can be taught and improved upon in the classroom, as well as in simulation exposures. Teamwork and communication can be emphasized in any training program. Situational awareness, in particular, can be enhanced through the use of cross-task cue utilization (Falkland & Wiggins, 2019).

Luxair (2013) states that the best way to be resilient is to be aware of the situations around you (and see Smith & Hancock, 1995). The training curriculum they suggest condenses the process to regain or maintain resilience into the acronym CABL$_3$. The idea here is to teach pilots to maintain vigilance and when they observe or predict a change to the risk profile at hand, they activate their "safety CABL$_3$." This process involves **C**ommunicating the observed change, **A**nticipating the consequence of the event, creating a **B**uffer or boundaries for behaviors, and then **L**ooking for the three critical indicators. These three critical indicators of decompensation, working at cross purposes, and outdated behaviors can be identified and remediated if the flightcrew is engaged in one of those behaviors (Luxair, 2013). This process is displayed in Figure 5.

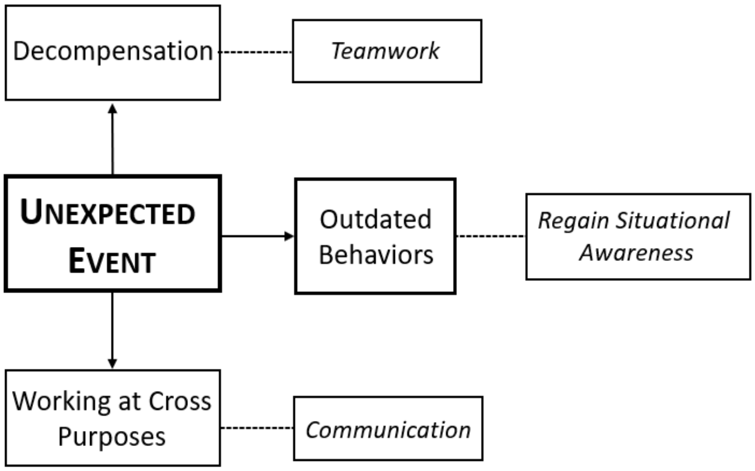

Figure 4. Illustration adapted from Luxair (2013) showing the incorrect responses to an unexpected threat, and the mitigating factors of each. (Dotted lines indicate each mitigating factor).

While the Luxair guidance discusses the $CABL_3$ model, a similar process is described in a 2015 report sponsored by EASA. In this process, three steps are suggested to recover from "mental upset," or startle effects. The three steps are:

1. "Unload," which refers to a technique for relaxing after being startled such as creating mental distance by deliberately pushing back into the seat, deep breathing, and conscious relaxing of muscles.

2. "Roll," which means regaining one's situational awareness which can be facilitated by each pilot stating out loud, "what they see, hear, feel, and smell."

3. "Power," which is technique to reinforce metacognitive skills by monitoring critical thinking and projecting the consequences of threats into the future by asking questions about the reliability and validity of the information sources and assumptions being used. (Field et al., 2015)

An evaluation of this methodology with 19 pilots completing the training showed positive effects for reacting to and managing unexpected events. The "unload" step appeared to improve pilots' information management and gathering of information regarding the unexpected event presented. The researchers note that self-efficacy can mitigate the effects of fear-potentiated startle and one of the effects of this training was to improve the participants' self-efficacy. The participants indicated via self-report, that they planned to use the training in future situations. However, in a follow-up with the participants, it was found that only a small percentage used the technique when faced with an unexpected situation in the operational setting. This may be a result of the shortened on-time training event, other crew members involved in the situation not having the training, and/or not having the technique as a standard operating procedure. As with other studies measuring resilience and responses to unexpected events (Landman et al., 2017), these researchers recommend the training be a part of the regular training cycle.

11. Measuring Resilience

Holbrook and colleagues (2019) have stated that there is no presently established way to measure resilience. This assertion was subsequently confirmed by Chandra and colleagues (2020), who agreed that there was, at present, no standard metric for resilience assessment and measurement (but see Hoffman & Hancock, 2017). Despite some degree of contention over the theoretical capacity to provide a direct, ratio scale measurement of resilience, it is generally agreed that resilience can only be measured indirectly, via the recording of observable behaviors. Chandra and colleagues (2020) found some examples of situations where pilots achieved positive outcomes through resilient behaviors, thereby avoiding potential mishaps. While these real-world examples are good ways to endeavor to determine resilience after-the-fact, it is necessary to consider the ways in which resilience can be measured without any risk to the aircraft crew or passengers, or those who might be affected on the ground. After all, a dangerous situation where resilience is required for survival, is not a good time for a pilot to determine that they are not, in fact, very resilient after all. For this reason, resilience must be measured *a priori*, before any adverse situation in which it will be needed. Thus, the need for the present foundational assessment.

Several studies have demonstrated the effectiveness of measuring constructs of resilience. For example, Gillespie, Chabover, and Grimbeek (2016) designed an instrument for OR nurses measuring perceived competence, self-efficacy, collaboration, control, hope, coping, age, education, years of employment as an OR nurse, and experience. A regression model was used to determine how the above constructs predicted nurses to overcome workplace challenges. The results of the study showed that hope, self-efficacy, control, and competence were statistically significant at predicting how nurses will overcome workplace challenges, whereas age, experience, years of employment, and education did not predict how nurses overcome workplace challenges. In addition, Prati, Pietrantoni, and Cicognani (2010) examined how self-efficacy contributes to the relationship between how rescue workers appraise stressful situations and their quality of life. The results of the study indicate that higher self-efficacy can lessen the impact that stress has on rescuers' perceived quality of life. "The study confirmed the hypothesis that self-efficacy might have a "recharge" effect. It was shown that rescuers' belief that they had a certain control in stressful events promotes resilience. The extent to which rescuers feel that they can face the challenges coming from their activities may function as a self-fulfilling prophecy. Expectations regarding adaptation lead to resilience. The results suggest that interventions aimed at developing rescuers' psychosocial skills are useful" Prati et al. (2010). Finally, the resilience scale by Connor and Davidson (2003) shows high psychometric properties such as validity and test-retest reliability and measures constructs previously mentioned by other scales such as self-efficacy, control, hope, competence, adaptability, perceived stress, experience and then new constructs not previously mentioned such as tolerance, patience, sense of humor, and commitment.

Hoffman and Hancock (2017) have proposed a time-based measure for assessing the ambient degree of resilience. They argued that such resilience measures could even be implemented in the short term. They also stated that in order to measure resilience and adaptivity, one would be more than well advised to use experimental manipulation. The use of such experimental manipulations would involve the immersion of a work system in several preplanned scenarios designed to prompt adaptive and/or resilient responses as mentioned in the variables from the studies above. This would allow investigators to test for resilience in pilots in low-stakes situation, without endangering human lives (Hoffman & Hancock, 2017). However, Matthews et al. (2017) warn that resilience in one situation does not always necessarily carry over to other situations and that scales may not be valid in every context. If this is true, then any measure of resilience which finds a pilot high in that trait may be accurate only within certain contexts. It is the agenda for future research efforts to specify the boundaries of these respective constraints, and the degree to which resilience as an individual or collective capacity, is transferable beyond the circumstances in which it is experimentally demonstrable.

12. Crew Resource Management

CRM training should, in theory, help to decrease the effects of both surprise and startle, by increasing the availability of emotional and cognitive management (EASA, 2017b). CRM should also manage the resources such as personnel, technology, and systems, develop appropriate use of behavioral responses, and help flightcrew to recognize any sign of degraded situational awareness while helping them regain their control. Flightcrew should have input in the details of a training initiative (EASA, 2017b). EASA also notes that those involved in the actual process of operating an aircraft should be in charge of evaluating training efficacy and determining how best to assess each competency. Additionally, EASA emphasizes the importance of training for automation use. Pilots should be able to use and monitor any relevant automation, especially during transitions as this can be a common period of time for misunderstandings. Those who will use any automation should be aware which level of automation is suitable for each task and context, which is usually provided by the manufacturers of the technology. They can also rely on Standard Operating Procedures to determine how autonomous a system should be during each phase of flight. This will prevent operators from relying on too high a level of automation and thus risking becoming removed from the system as a whole.

CRM training methods are varied. Respondents of a survey indicated that more time should be allocated to CRM training, and new assessment forms should be developed. Often, the information upon which new training interventions are based come from sources such as operational risks like accident reports, news, and safety analyses; training issues such as confusion regarding competencies or failures of assessments; and compliance issues such as not properly using new systems (EASA 2017b). Wing and colleagues (2020) posit that while degradation of pilots' manual skills due to increased automation is a problem when humans are left out-of-the-loop, there are some benefits to this situation. Training for pilots becomes cheaper and easier, which allows companies to focus more on training other skills, such as resilience, rather than focusing entirely on technical abilities. They do maintain that a human is necessary in the flight deck, for cybersecurity reasons at the very least (Wing et al., 2020).

Hancock and Cruit (2020) have discussed the concern that no purposed-directed empirical evidence exists to confirm the effectiveness of EASA's resilience training within CRM (EASA, 2017a). They propose that, as the mandate for resilience training was initially based on training for cognitive flexibility and adaptive performance, it does not focus on 'resilience' in the distinct and definitive sense of the term. They observe that expertise is not only a measure of training or, more specifically, not of training alone. This assertion is supported by evidence from Cruit and colleagues (2019) in a study where they found that among a group of pilots who successfully avoided disastrous consequences during an unexpected situation, flight hours and classroom training was not the sole predictor of resilience. They go on to state that those with high confidence in their own technical skills have an easier time in considering all possible courses of action in difficult situations, as compared to those with low confidence (Cruit et al., 2019). While one may argue that confidence arises from expertise, it cannot be ignored that these data indicate that confidence leads to greater success in problem resolution. Quite naturally, unexpected physical events that occur in flight will strike differing flightcrews in differing ways. The purpose of a safety-oriented flight training system is to ensure that, even in the face of these trait variations, successful resolution is assured; to the degree that such an indemnification is possible in an uncertain world.

CRM can be used to maintain safety because according to Faerevaag and colleagues (2018), the goal of CRM is to reduce the occurrence of, and magnitude of, human error. Additionally, operational risk management (ORM) can maintain safety through on-going, dynamic, and rectifiable risk assessment. When flightcrews use the five step ORM process, they can make better risk-related decisions, avoid hazards, and maintain safety. The steps of this process are to:

- Identify the Hazards
- Assess the Hazards
- Make Decisions
- Take Control
- Supervise and Monitor

ORM is quite similar to Risk and Resource Management (RRM), the steps of which are to:

- Assess the Situation
- Determine Resources
- Communicate
- Act
- Debrief

Each of these sequences are essentially recapitulation of the human information processing system (and see Broadbent, 1959). What they require is the ability to engage in sensory assimilation of the immediate circumstances and this distinction of the ambient sensory pattern is encapsulated in the initial stages of: *Identify the Hazard, Assess the Situation*, as noted above. These latter steps, of course, is fully bound up in the whole realm of situation awareness, for which there is a plethora of extant research work (Endsley, 1995; Smith & Hancock, 1995). Information assimilation and associated perception are necessary components of problem resolution, but in and of themselves are insufficient for solution. Thus, the next two, comparable stages: *Assess the Hazard, Determine Resources*, represent slightly differing processes. Hazard assessment is an act of perception, resource determination is a preparatory stage to problem-resolution. Nevertheless, each fit along the path of the information processing sequence. Both of the next two comparable steps: *Make Decisions; Communicate*, represent the processes of decision-making and their externalization to other members of the airspace operations system. Decision-making and its potential modification from external sources of information (e.g., ATC input), represent a key step in responding to unexpected events. Much here, as we shall emphasize later, is contingent upon both the actual time available for resolution and, critically, the *perceived time* available for resolution. These two respective durations are not necessarily, and even rarely of the same magnitude. The following steps, *Take Control, Act*, we consider synonymous and this represents the information processing stages of response selection and response execution. As is clear in many adverse outcomes in aviation, these two components of action are themselves, not always mapped directly on to one another. Correctly identifying the appropriate response to enact does not always mean that the physical action (e.g., an accompanying button press) is always carried out as planned. Interestingly, in aviation as it is in other technologically mediated domains, it is possible to witness the paradox of two wrongs making a right. That is, an operator can make an incorrect decision to respond, but then make a subsequent error in response execution which paradoxically leads to success. The influences of stress on these respective stages of response are vital and have to be emphasized in training. The final, *Supervise and Monitor, then Debrief* pair are part of the cybernetic feedback loop that requires us to monitor the results of our actions in order that subsequent learning, which is a component of resilience, can occur. Given the rarity of actual unexpected events in flight, training requires that we learn, in a meta-level sense, from the prior events that have been witnessed, and the prospective events that we might envisage as occurring. These latter, 'envisioned world' approaches represent one of the greatest challenges facing the future of safe and effective aviation operations.

Faerevaag et al., (2018) observed indoctrination procedures and noted an emphasis on safety during training. The focus of all operational decisions tends to be on maintaining safety according to the air carrier's principles. Threat and error management (TEM) represent an extension of the generation of CRM which aimed at eliminating, or at least substantially reducing, errors by flightcrew (Martin, 2019). This initial goal was quickly changed as the impossibility of eliminating all errors became clear. Therefore TEM, which advocates error tolerance, became the next movement (and see Hollnagel, 2017). TEM involves managing the errors that do occur. By such actions, the aspiration is to mitigate risk from both expected and unexpected events (Helmreich et al., 1999). In regard to TEM, pilots are aware of potential risks and have plans for those events (Martin, 2019). This concept therefore shares some similarities to resilience and its training. However, resilience is more widely applicable across these and other systemic challenges.

According to Faerevaag (2018), the TEM model, as used in the training of flightcrews, teaches them how to prevent errors with the resources they have on hand, how to reduce the outcomes of errors, and how to forestall undesirable aircraft states. All these factors can help to prevent an incident. However, as valuable as these resources are, there is always a question of whether the training, however it is structured, truly transfers to practice in the moment. Since CRM performance cannot be directly measured by any one metric, it can only really be assessed using behaviors that are observable. Flightcrew communication skills are one factor that can be observed. The ability to manage workload is another critical skill that may be learned through CRM. Planning and decision-making go hand-in-hand, and an individual's skill in either of these dimensions reflects upon the degree of their effective training. Leadership, and the ability to work well in a team are also important skills, as is the ability to maintain command when needed. The monitoring and management of non-human resources is another critical, and observable skill (FAA, 2004). However, while many CRM training courses cover generally similar topics, training regimens do differ by air carrier, and each carrier tends to focus on some aspect of CRM behavior based on their company's culture (Faerevaag et al., 2018).

As critical as these aspects are, there is relatively little research on the relationship between observable behaviors and CRM elements (Faerevaag et al., 2018). Therefore, it is difficult to determine whether CRM training is effective or if it needs to be altered or amended in some fashion. EASA has published a list of recommended topics to cover in CRM training (EASA, 2017a). This list includes resilience, how to deal with being startled, and how to deal with being surprised. However, other international aviation bodies such as the International Civil Aviation Organization (ICAO), UK's Civil Aviation Authority (CAA), and International Air Transport Association (IATA) all have their own requirements and their own lists of recommendations for what should be part of CRM training. According to Faerevaag et al. (2018), stress management, fatigue, interpersonal relationships, and similar topics shown to influence individual resilience should also be included. According to Martin (2019) CRM skills are key components of team and individual resilience and are a way to improve pilot responses in adverse situations. He draws an analogy to the steps in TEM where crews avoid, mitigate, and manage threats and errors. Strategic resilience can be improved through anticipation of unexpected events and good technological knowledge which will help crews avoid the threats and errors. Tactical resilience is the mitigate and manage steps of TEM. It can be improved through rehearsals of responses to critical events, understanding of associated cues, and emotion management during critical events (Martin, 2019)

13. Proposed Definitions from this Research Assessment

In light of above observations, we begin our present review by looking to establish the critical meaning of terms that are used in the field. Here, we have not attempted to provide "perfect" definitions of each of these key terms. Rather, the purpose of this section is to ensure that (a) our team can express and communicate a mutual understanding of the constructs we are researching, and (b) we provide a definition and/or description of the constructs that is clear to the readers of this report. We need a common foundation and realize the many different connotations that many of our study terms convey (Hancock & Volante, 2020).

With this goal in mind, we have compiled the definitions and features, which have been selected from published research and reports, of the constructs pertinent to safe reactions to unexpected events. Each of the relevant construct has our "working definition" which is associated with a brief explanation of why it has been either created or adopted by the present research team. These respective definitions are then referenced to a series of associated tables that provide a much more detailed overview of each of the individual concepts to hand (see Appendix A).

It should be emphasized that the constructs described within this section are all interrelated. For example, resilience is a construct that is comprised of many different terms such as its overlap with adaptability, cognitive flexibility, performance adaptation, metacognition, etc. Thus, some of the information conveyed in relation to each term necessarily overlap each other. In two cases, we present charts in order to compare and contrast closely allied constructs that are particularly intertwined, that of resilience and adaptability, and startle and surprise. The literature search used to create our definitions is not exhaustive. However, we have included major historical and modern research studies expressed in the context of the purpose of our present study project. The full citations for the presented information can be found in the Reference section at the end of this review.

13.1 Resilience and Resiliency

There have been many definitions of resilience. A selection of definitions is presented in Appendix A. While the definitions and uses of resiliency may vary in terms of specific phrasing, there are key similarities which all point to resiliency as a factor in recovery from adverse events and outside threats. Our working definition of resilience with focus on the system is *the systemic capacity to change as a result of circumstances that push the system beyond the boundaries of its competence envelope. The system may have to amend some, or even all of its goals, procedures, resources, roles, or responsibilities. As a result of those changes, the work system then expresses a revised competence envelope. In effect, it becomes a different system.*

We consider resilience to be the capacity of a system; it is not the system itself, but the form of behavior that it exhibits. Here the term ‘system’ is intentionally interpreted widely, although our particular concern in the current project is with the evolving complexities of airspace operations. One predicate of resilience is an *a priori* stable state of behavior. This does *not* mean a system which is fixed and inflexible. And in fact, we anticipate it being likely that resiliency will covary with highly dynamic, adaptable systems. Such systems are under constant states of perturbation from sources both intrinsic and extrinsic to that system.

We also consider a definition of resilience that pertains directly to the human in the system which is, *the ability to adapt to changing circumstances by attaining a differing form of operational stability through situation assessment, self-review, decision, and action.*

While the definitions vary, all describe resiliency as a personal characteristic which is integral to overcoming adversity. It involves the ability to take in new, sometimes very unexpected, information and adjust actions as needed to survive and manage the adverse event. Those who have prepared in the past are ready for a specific practiced situation. Those who are resilient are capable of determining how to act in a completely unique situation.

There are many concepts which are similar to resiliency. Adaptability, and cognitive flexibility, are both similar. However, there are some important differences between resiliency and adaptability. The main difference is the end goal; the goal of adaptation is to return to its original state, the goal of resilience is to thrive in a “new normal” (Hancock & Cruit, 2020). Therefore, resilience embraces the challenge of functioning in novel circumstances, mitigating risk that is, perhaps, unknown.

13.2 Adaptiveness (Adaptability, Adaptive Expertise)

Adaptability and resilience are not mutually exclusive terms, and often the resilient are adaptive and vice versa. Chandra and colleagues (2020) call adaptive experts flexible and innovative, noting that they will be more capable of applying their knowledge to unexpected events; a fact which makes them better able to manage stress and reduce errors.

Cañas and colleagues (2005) call mental flexibility an ability, which allows people to use cognitive strategies from past experiences and apply those same strategies to new conditions. This adaptability will help in the same sort of situations where resilience will be useful. While resilience may prove to be more necessary in conditions which cannot be altered, adaptability is a suitable correlate. However, as valuable as they are, adaptability and other forms of mental flexibility may not be easy to train. Faerevaag and colleagues (2018) clarify that pilots must be ready to react to unanticipated situations, for which they have not been trained. To attempt to train the skill of mental flexibility is nearly impossible as the situation in question *cannot* be one with a known procedure. If a particular threat is known, it would be more beneficial to train pilots to respond to it correctly, than to train them to respond in a flexible manner. Additionally, mental flexibility can diminish performance in some cases when a pilot, operating under new procedures, must return to standard procedures. Overall, more research on the topic of mental flexibility is needed to determine its value. There are many aspects of resilience and one of the major components is related to that of adaptiveness or adaptability. Experts know what *to do* as well as what *not* to do in a situation. In addition, *adaptive experts* not only use what they know, they monitor their current level of understanding of a situation, continue to learn, and strive to move to a higher level of functioning. They use each new situation as a challenge which provides a forum to facilitate additional expertise (Kochan, 2005).

13.3 Comparing Resilience and Adaptability

Given the strong interrelationship between resilience and adaptability, we collected information that shows the similarities and differences of these terms. Largely, definitions of adaptability focused on an ability to return to an ideal state, while resilience definitions emphasized the ability to continue operations in a new state. We here use Hancock and Hoffman's (2017) contrast of the two constructs: "In a general sense, we can view adaptability as the search for stability in an already occupied parameter space. Resilience in contrast is the achievement of a new state of stability in a different parameter space."

13.4 Cognitive Flexibility

According to Spiro (1995), "*cognitive flexibility is the ability to spontaneously restructure one's knowledge in many ways, in adaptive response to radically changing situational demands."* This has been adopted as our present working definition. That is, in complex environments, learners generally cannot retrieve a completely intact and applicable learning structure from memory alone. Instead they must combine, recombine, and reinvent structural components to meet the requirements of each particular situation (Spiro, 1988). Thus, specific, and useful operational knowledge becomes, to an important degree, context dependent. In general, cognitive flexibility theory has been associated with the world of pedagogy, involved with both methods of teaching and modes of learning such that the individual can assemble a flexible and useful lexicon of world knowledge.

Traditional modes of learning provide fact-based instruction around which themes, threads, and patterns are woven and extracted in order that higher-level abstraction-based thinking can be engaged. In traditional approaches to aviation training, the emphasis is on procedures and such traditional instruction techniques are largely predicated upon the assumption that aviation is largely a closed-end world of knowledge. This is not to say that aviation, and more especially exploratory elements of aviation such as airspace operations, cannot be obscure and complex. Assuredly they are. However, the fundamental physics of flight are relatively well-known, especially for regions in which commercial aviation predominantly operates. Thus, the critical outline of the operational context is known and relatively fixed. With that foundational assumption, the instructional mode which is predicated upon this basis, seeks to enact, if not algorithmic solutions, then at least algorithmic search for any sources of disturbance and failure.

As is clear from the excellent record of safety, relative to other transportation segments and other safety-critical complex systems, this approach has to date worked reasonably effectively. However, we are now entering new realms of operation in which the driving forces of demand and technology add layers of additional complexity that serve to defeat immediately searchable lexicons of resolution algorithms. These circumstances mandate that training regimens now begin to explore response abilities expressed in concepts such as cognitive flexibility in order to support resilient response to unexpected and unanticipated events.

13.5 Metacognition

Our working definition of metacognition for the present review is: "*the ability to monitor one's current level of understanding and decide when it is and when it is not adequate*." This definition is derived from Kochan (2005). More commonly, metacognition is the awareness of one's own awareness. It represents a form of personal knowledge and is, at the same time, a skill which can be used to control and manipulate differing cognitive processes (Wells, 2002). It is thinking about one's thinking, but in the present context it is also a positive and elaborative process. It is not simply a form of daydreaming to no purpose no end. This is because the fostering of meta-cognitive skills can be central to evaluating one's own line of response to some incipient emergency or unanticipated demand (Magno, 2010). There is, of course, an insidious trade-off in respect to meta-cognitive activity, most especially in emergency situations. That is, can one afford the time to engage in this level of self- evaluation while precious seconds are passing which may prove determinative. Traditional safety approaches tend to be very retrospective and judgmental in their determinations. Thus, avenues of solution and remediation are, almost necessarily and inevitably, evident in retrospect. However, in prospect, these solutions are often observed at best, obscure, or even totally 'hidden' at worst. Metacognition enters this fray when a line of response has been adopted and the pilot(s) must retain a skeptical attitude to the specific course that they have adopted. Resilient response involves problem-solving on-the-fly, as it were, and commitment to one single form of potential resolution can act to 'lock-out' other lines of potential progress. This is especially so because we know that attention becomes focus and narrowed, most particularly under extremes of existential threat (Hancock, 2021a). In light of these 'focusing' effects, it may be here that 'team resilience' becomes most important. For, while one individual can be involved with managing and effecting a potential line of resolution, e.g., the pilot in command, other members of the team can still be engaged in resolution path search. Prospectively, one of this immediate response team will be the automation and autonomy instances to the aircraft itself. Here, we have another looming issue that must be explicated and faced.

Much has been made of the coming transition of technology and its transformation from "tool to team-mate." Such ideas are all well and good. However, in an emergency situation, it will be this very 'team-mate' who is presenting the information concerning the unexpected events and, whom may be identified as the 'problem' to be solved. Thus, the automation is the supposed team-member who will appear to be malfunctioning. The long-term solution to this issue involves the creation of certain companion or guardian entities, whose task is to act as the 'metacognitive' property of the supporting computer systems at hand. As yet, these 'shadow systems' have yet to be fully and operationally developed. In the same way that CRM has been successful, we now have to introduce ARM (automation resource management). It might be noted that this is already the function of the human flightcrew, but we need a much more extensive and explicit focus upon this collaboration. The interfaces should also provide a separate configuration and functionality from that of its primary flight necessities. In short, we need metacognition on the flight-deck, and in other segments of advanced aviation. However, those metacognitive capacities need to extend to all operational entities and not be confined solely to the systems' human members.

13.6 Novelty

We here have adopted the definition of novelty as: "*a property of a stimulus that has not been previously presented to or observed by and is thus unfamiliar to the subject*" (Gordon & Luo, 2011). There are many circumstances in the world which bring unique arrays of stimulus features to an individual. Obviously, the baseline of novelty covaries with age and maturity such that there is necessarily a greater number of new things in the world to those who are new in it. However, our focus here is narrowed to the situation of aviation operations. The definition adopted is somewhat constraining. This is because the authors are looking at novelty at the stimulus level. At this granularity, it is unlikely that an experienced pilot will encounter a specific stimulus or even range of intensity of that stimulus, that they have not experienced before. Thus, we would elaborate upon this definition to define novelty in terms of originality in the individual's perceptual field. This allows us to parse novelty into the interactive components of an array of stimuli. Since our concern is for unexpected events, the degree to which expectations, informational value, and novelty covary, proves to be of central importance. Thus, are all unexpected equally novel? Does the degree of information intrinsic to the previously unobserved perceptual display provide an index of such novelty? And most especially, how does the degree of novelty presented influence the possibility of a successful action. In part, we must explore whether all novelty is equal. That is, although an individual may not have registered a particular perceptual display before, does its similarity with experiences that they have had then impact the inherent information present and the opportunity to explore and achieve successful resolution. This is a critical element of training for resilience for unexpected events and one we return to in more detail in what follows.

13.7 Startle

For the purposes of this study, we define startle as, "*An uncontrollable, automatic muscle reflex, raised heart rate, blood pressure, etc., elicited by exposure to a sudden, intense event that violates a pilot's expectations"* (FAA, 2017). In general, this covers the issue and emphasizes that startle is largely a transient event. It is, as we report, contingent upon the rise time, the intensity, and the duration of the startling stimulus. In some senses, startle is a relational concept since it implies a degree of non-expectation. Yet, even given this relational element, the forms of startle that are encountered in aviation are not the sort of events for which training is an effective solution. Indeed, it might well be argued that startle effects might even prove helpful to performance, although the term is most often associated with adverse events and even consequences. Startle in aviation can derive from all five senses, but as with other forms of transport, vision and audition tend to dominate. Not least because these are the most utilized forms of sensory input via the formal interfaces designed for control. We here draw a distinction between *sensory startle*, which is that most commonly envisaged when the term is employed, and *cognitive startle*. The latter, cognitive dimension, may be associated with intensity of sensory experience, but the connection is not a necessary one. Indeed, cognitive startle can derive from a series of sensory inputs and on-going mental model of operations which is not associated in any way with sensory intensity, rise time, or duration. Cognitive startle appears to be much more associated with the fracture of expectation, as evidenced in circumstances characterized as mode error awareness (e.g., Sarter & Woods, 1995).

A startle effect occurs with a sudden onset change in the sensory environment and can be observed in any of the human sensory systems. The principal difference between each of these sensory modes is with latency of reaction and of course, the issues involved in trying to ensure comparative psychophysiological equivalency across these systems. The information provided in this brief section derives mostly from startle induced by acoustic change although, in principle, the observations apply to all. One of the primary drivers of startle is stimulus ramp time. This can be of any duration but with the human auditory system, any ramp time below 2ms would be considered ‘instantaneous.’ Startle reaction strength covaries to some degree, with extensions to this ramp time, with an asymptote of the effect occurring around 50ms which tends to bound the duration of the ‘instantaneous moment’ (Poppel, 1988). The ramp time effect is naturally crossed with stimulus intensity. The frequency of the startle effect increasing up to an asymptotic threshold with extreme levels of intensity. Whether the precise degree of startle is locked to these physical properties is a matter of empirical exploration. However, we are primarily focused on the practical realm of aviation operations, induced startle over sensory mediated thresholds are considered sufficient for inclusion. Startle effects that cannot be habituated to, can also induce acute and chronic stress effects that lead to health consequences. However, and again, we are dealing in aviation with evidently rare events and so these influences are not considered further here (but see Szalma & Hancock, 2011). In summary, the primary physical drivers of startle intensity are stimulus rise, time, peak intensity, and stimulus duration, each serving to increase the startle reaction.

It may well be that in the realm of complex operations, the word startle carries with it rather negative connotations. However, when viewed in evolutionary terms, the reaction proves rather an important and positive one (Koch, 1999). Startle effects serve to induce a state of heighten arousal and a readiness to act and respond to the challenges and issues that the startling stimulus presents. This arousing and protective response is, phylogenetically speaking, a low one and appears in many orders of living systems. The advantage that it confers is also confirmed by the cross-wiring of startle effects expressed in differing sensory regions of the brain (Yeomans et al., 2002). It is probable that such brain architecture plays a role in the facilitation effect that startle stimuli appear to exert upon responsivity represented in measure of reaction time. In part, the degree of cognitive distraction, often perceived to be associated with startle, can be ameliorated by a pre-potentiating stimulus. But the application of the latter requires the experimenter to know and control the ‘state-of-the-world.’ This can be easily achieved in the Laboratory but is very rarely possible in real-world conditions. Indeed, if one knew of the circumstances that would induce a startle effect in aviation, for example, it would already be addressed by some relevant and anticipatory technological response and thus would not be startling. In sum, startle effects are of some degree of concern in the etiology and progression of many general ‘unanticipated events.’ However, their actual observed impact appears, from epidemiological forms of survey to be small at best. As a result, we do not here recommend any form of training that is predicated upon these specific and transient responses. While it is important to be cognizant of their influences, the more general concern for training efficacy most probably needs to lie with other dimensions of operator capability.

13.8 Surprise

We take our definition of surprise in aviation here to be: “*an unexpected event that violates a pilot’s expectations and can affect the mental processes used to respond to the event”* (FAA, 2015). In formal terms, surprise provides a high degree of information gain. This is because in terms of both its spatial and temporal characteristics, surprise events tend to maximize uncertainty. In purely formal terms, the general term 'surprise' must connote some perceptual threshold (Meyer et al., 1997) which, presumably, covaries with the exposed individual. For example, a 'surprise' birthday party is, to a degree, contingent upon the individual's knowledge that it is their own birthday. In this way, we can see that context and personal knowledge each contribute to the severity of the surprise event. In safety critical systems such as advanced aviation, we seek to minimize surprise in all its forms. Since the degree of surprise, and then assumedly its down-stream effects, covaries with the knowledge of the person to hand, training can have a substantial influence in mitigating surprise level. As with virtually all analog traces, there is a time-course to surprise events and we can look to this general understanding to guide us here.

As mentioned, the definition of surprise that we employ here is defined as: "*an unexpected event that violates a pilot's expectations and can affect the mental processes used to respond to the event."* (FAA, 2015). We have selected this particular definition due to its relevance to aviation operations. Surprise itself is directly related to formal information theory. That is, information is conveyed when uncertainty is reduced. Surprise represents a circumstance in which there is a significant degree of uncertainty. In actuality, surprise occurs with a sudden fracture or brisance of the currently operating metal model of a situation. In consequence, surprise is almost necessarily constrained to occur in situations where the person or operator has already reached an a priori degree of certainty about which course of events will occur. In consequence, 'surprise' itself is simply a threshold state on a spectrum of information foraging (Pirolli & Card, 1999). We can interpret surprise within this foraging concept as one in which the bottom-up sensory input, which in strongly procedural environments often serves simply as periodic confirmation of expectation, now reverses that function and indicates a state of events directly counter to that which is the one that is anticipated. In some ways, these sudden inversions are similar to the point at which differing interpretations of ambiguous figures, such as the Necker Cube, reverse themselves. In other sub-disciplines of behavioral science, these moments of revelation and state reversal have been modelled using 'cusp-catastrophes' derived from theories associated with non-linear dynamics (and see Guastello, 2013).

In general terms there are two major characteristics of a transient analog signal and they are onset rate and amplitude. As implied by what has been observed, surprise has to be represented by a critical combination of these two aspects. In terms of the temporal profile, the onset of the event itself can be a sudden one. This might be typified by a mid-air collision when two aircraft come into direct contact. Such an example might be seen in an event such as the Proteus Airlines Flight 706 collision over Quiberon Bay in northern France (Bureau Enquêtes-Accidents, 1998). Here, the flightcrew had little to no time to recognize an event, which is often considered 'instantaneous.' Conversely, adverse conditions may build up slowly. The systemic failure involved may either have below threshold recognizability at first onset or may progress at a rate that is also below threshold for a period of time. Here, the instant of surprise is not solely a property of the operational condition, but the flightcrews' recognition of it. The latter itself can be a short duration 'instant of recognition,' despite the fact that conditions underlying that surprise have been building up for some period.

As with onset time, the degree of surprise is also dependent upon the degree of disturbance experienced. Certain events can be surprising but have little practical consequence. Others can be of catastrophic significance. One of the important features of understanding human factors facets of event accounts is to know that the same basic forms of information processing capacity are involved in surprise response, whether the implications of that event are trivial or major. In the present work, while we are concerned with transient startle effects, which often follow upon surprising incidents, we are more focused upon the dimension of surprise. The latter two linked response patterns are compared and contrasted below.

13.9 Comparison Between Startle and Surprise

Kochan (2005) has explained the difference between startle and surprise by noting that startle is predominantly a physiological reflex which lasts in the order of milliseconds to seconds. In contrast, surprise proves to be a more prolonged, emotional reaction. Kochan (2005) concludes that: "*Surprise is an emotional response to an unexpected event. Example: The airplane stalls when on autopilot at high altitude owing to insufficient thrust available to maintain airspeed while keeping a level altitude, as the pilot had become distracted and was not able to monitor the flight properly whereas startle is a reflexive response to an unexpected event. Example: a pilot winces and blinks after a windshield suddenly cracks in flight*." Kinney and O'Hare (2019) noted that surprise and startle can both arise when the information that one is perceiving is in contrast to the active frame of thinking. They further asserted that "reframing" of the situation, through effort and reasoning, could mitigate any negative outcome. A conceptual model, created by Landman et al., (2017) shows increased and decreased effort during startle or surprise events (See Figure 6). Figure 6 begins with the triggering event that stimulates the senses. Depending on the human's perception of the event, it would lead the human to either appraise the stimuli fast or slow. A surprise is indicated when there is a mismatch between what the human expects and what is appraised. A startle response occurs when the human perceives a threat. Moving down the model, it shows how a selection and execution of actions can lead to a change in stimuli. As previously mentioned, and indicated in Figure 6, reframing the situation can change the stress response. The sequence of events and potential influences are represented by a solid and dashed line, respectively. Thresholds are indicated with a double line. Minus and plus signs indicate increased and impaired effort.

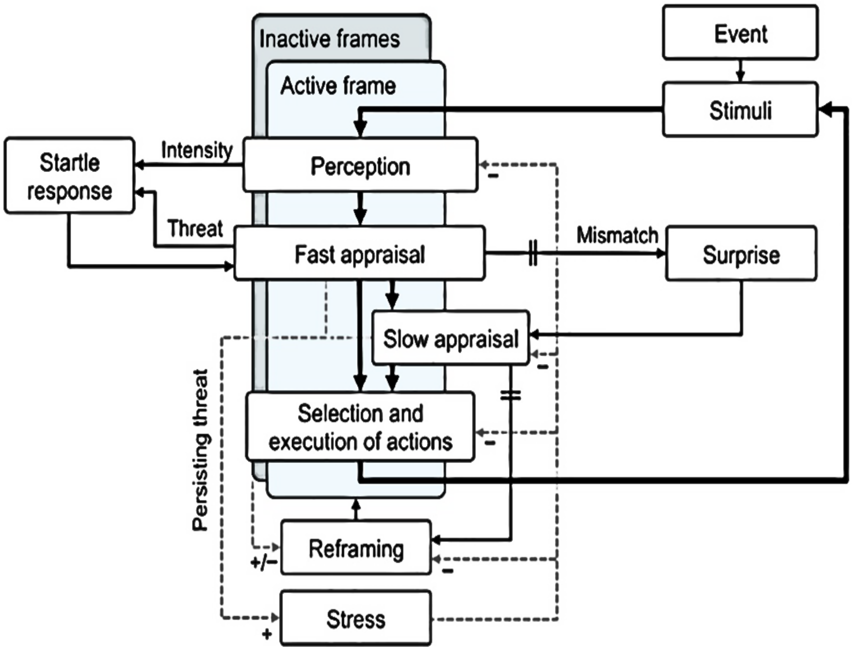

Figure 5. Landman et al., (2017), model of startle and surprise. Sequence of events and potential influences are represented by a solid and dashed line. Thresholds are indicated with a double line. Minus and plus signs indicate increased and impaired effort.

13.10 Stress

"*Stress is a response to threatening situations that involves biological, cognitive, behavioral, and emotional components*" is a general, aviation-based definition that we have adopted from Dismukes et al. (2015). Stress itself can be thought of as an external factor, arising from environmental conditions (e.g., heat, noise, etc.), and at the same time it can be conceived of as an appraised status by the exposed individual (Lazarus & Folkman, 1984). The latter, transactional perspective features the ability to incorporate the response of specific individuals into considerations (e.g., age, experienced), the former, physics-based descriptions, often used for occupational standards, tends to feature objective measures which are specifiable independent of any individual who may be exposed to them. The Dismukes et al. (2015) study falls largely into the transactional paradigm. What is needed, of course, is the specification of the level of stress and then how that perceived or actual challenge influences the individual pilot's capacity to respond. This is a common challenge in applied research which requires us to step from the general, nomothetic level of understanding to the individual, idiographic, and bespoken level of actual, real-world response, The vast improvements in computational capacity encourage us to believe that progress can be made along this axis of ever-greater specification (and see Hancock, Hancock, & Warm, 2009).

13.11 Safety

The working definition of safety for the present review, we derive from ICAO (2018); "the state in which risks associated with aviation activities, related to, or in direct support of the operation of aircraft, are reduced and controlled to an acceptable level" (p. vii). Of course, this is a rather contentious definition since safety is such a difficult and inherently ambiguous area of concern. Safety has been described as a “dynamic non-event,” but this approach has been somewhat criticized, especially because the unit of the ‘non-event’ cannot be specified, for example, on a ratio scale (and see Hancock et al., 2019). Our focus here, on unexpected events, also runs counter to the growing zeitgeist in safety which now looks to feature more focus on systems’ success rather than its infrequent failures. This is especially true for aviation which, in itself is almost necessarily a rather hazardous enterprise, but is one that has proven remarkably disaster-free across its century of effective commercial operations. This statement, of course, appears to be a polemic one, especially because the human mind almost inevitably focuses on events which did occur and remains almost oblivious of those that did not. The ICAO definition is certainly justified in referencing ‘acceptable levels’ and these are set by social consensus, strained through varying cultural and political processes evident in differing parts of the world. The challenge lies in two other elements of the adopted definition; these are risk and management. Risk assessment, at its very heart, proves to be an informed estimate of future conditions predicated upon past experience. As has been noted, the degree to which the future resembles the past, the greater the efficacy of that assessment. But critically, vice-versa. What can be assessed can be managed. What cannot be assessed leaves management in the dark. These are the challenges of prospect and the ‘envisioned world’ problem. Our indemnification here centers primarily on modelling and simulation of possible future conditions. Thus, appropriate training for safety for unexpected and unanticipated events is contingent upon the degree to which they can be imagined, modeled, simulated and the communicated to those likely to be exposed to them. What we do not presently know is the relative frequency of these unanticipated events and whether, in the future, they will increase, decrease or stay the same. Into this mix of safety enters new forms of control technology which bring along their own assessment and reliability challenges. It is preparing the human operator in these evolving systems which represents the focus of our present program of research.

13.12 Monitoring

Our working definition for monitoring is adopted from the FAA’s Crew Resource Management Advisory Circular (FAA, 2004). This definition states monitoring is: “the observation of the aircraft’s flight path and systems and actively cross-checking the actions of other crewmembers.” As far as it goes, this represents a useful foundation. Yet what is required to further explicate monitoring activities is the processes by which they prove to be successful or unsuccessful and the context in which those successes or failures are experienced. Largely, this brings us to the challenges of vigilance (Hancock, 2013; 2019). Here, individual or multiple displays are presented to an observer and the imperative laid upon them is to watch for critical deviations of differing threshold values. When the display is a singular one, when the signal itself is at near threshold, when it is embedded in a sequence of non-signals, and when it is presented under stressful conditions, the individual tends to diminish in their response capacity. This diminution of efficiency is known generally as the ‘vigilance document.’ Much time and effort has been given over to researching this decrement and searching for effective ways to counteract its presence (and see Hancock, 2013, 2017). Perhaps the most important dimension here is the design of the work-task and the design of the interface to that work task. For, if these are designed effectively, much of the decrement can be mitigated. However, it remains the fact that human beings are generally poor at responding to these forms of task and it is advisable to try to circumvent their presence via design, and not by training. Although it is possible to train for vigilance, this is a remedial and unadvised strategy. Reacting to unanticipated events, as we have seen, requires that they be registered in both sensation and perception. Vigilance-type demands often mean that, even if they are not actually missed, signals still take longer to process, and that increased latency may well be vital in a mission-critical and time limited circumstance. Such observations confirm that some of the issues facing members of the advanced aerospace system are themselves systems-level challenges and therefore have to take on at those differing respective levels of concern.

13.13 Robustness

In respect of the dimension of robustness, we have chosen to use the definition from Hoffman and Hancock (2017) who defined it as: "*the ability of the system to operate within its normal operating boundaries when it is perturbed*." This represents an expansive category that is linked to adaptation, upon which we also comment in the present sequence. Robustness is then a property of extent. That is, the greater the range of circumstances over which a system can maintain operational stability, the greater the degree of robustness that it presents. In general, and as is also communicated in the general parlance usage of the term, robustness is a good thing. It implies the capacity to resist perturbation from whatever source. In terms of the operational ream of the system to hand, we can expect that robustness is spatially distributed such that maximal robustness is expressed at the point of greatest (central) stability. As the system is required to venture into more remote regions of its operational capacity, the degree of robustness it expresses is anticipated to diminish accordingly. Specifying exactly what measures of the system characterize robustness, and exactly what challenges in the environment serve to diminish (or perhaps even enhance) this expressed capacity is a general challenge. The aviation system, being a well-studied and reasonably well comprehended one, would seem to be a prime candidate to leading the specification and measurement of robustness. Exactly how the properties of robustness and resilience overlap becomes a challenge for our present sequence of evaluations into reactions to unexpected events. It may well be anticipated that a system that is characterized by a high degree of robustness can tolerate greater degrees of unexpectancy, although this has yet to receive specific experimental evaluation.

13.14 Brittleness

Essentially, brittleness is the antithesis of resilience. From a human perspective, **metacognition** and other human cognitive abilities have the potential for overcoming the opacity and brittleness of traditional systems. As applied to technology, resilience to the system can be improved by permitting computational agents access to the reasons for decisions and actions. Including these aspects in their programming allows autonomous changes in decisions, thus increasing the flexibility of the human-machine system. It is useful to look at the computational underpinnings of system brittleness as it becomes useful in the design of more resilient human-technology teams such as those in the aviation domain.

David Woods (2015) defined of brittleness as "*a rapid fall off or collapse of performance that occurs when events push a system beyond its boundaries for handling changing disturbances and variations.*" This definition links directly to the concept of resilience and indeed, how it has been proposed to be measured (Hoffman & Hancock, 2017). Interestingly, both can be indexed by the allied notion of hysteresis (and see Jansen et al., 2016; Morgan & Hancock, 2011). The latter value indexes the degree of recovery of a system back to its original state. The implication of brittleness is that the system is vulnerable to a complete breakdown. Presuming here that the greater the degree of associated brittleness, the more the associated risk of such catastrophic collapse. In more formal terms the hysteretic value is close to zero. Resilience also implies a low hysteretic value here, with the exception that the eventuality is not total collapse, but the securement of a new stable state of operations. High degrees of adaptability imply high values for any hysteretic constant (and see Hoffman & Hancock, 2017). The fundamental difference lies firstly in the mode of transition observed, and then secondly the functional outcome of that transition.

The general architecture of these transitions, most especially with respect to individual human performers under the driving forces of extreme stress, have been examined by Hancock (2009). In respect of general trends, what we see here is a drift of the first moment of the performance distribution; namely the mean levels of efficiency as commonly expressed by a combination of speed and accuracy. These drifts do not necessarily have to reflect ubiquitous downward trends but can be exhibited by systematic changes over time; of which more below. In many ways, these are akin to the information provided in Statistical Process Control (SPC), and if the appropriate metrics for large-scale systems can be distilled, then these self-same forms of data can and would be useful in monitoring aviation operations. As with all such proposals, the challenge is in securing the appropriate mix of relevant measures. However, changes in the first moment of the distribution (e.g., mean, median, mode, etc.), are necessary but not sufficient for monitoring system status. What is especially important here are representations of variability across time. These reflections can be seen in changes in the second moment of the response distribution, and in the behavioral sciences are tracked in terms of standard deviation, standard error, variance etc. It is important to note that differing performance assessment disciplines tend to use differing reflections of this variation (e.g., RMSE and steering entropy for driving, etc.). In and of itself, variability measures are important and provide insight into the revealed dynamics of behavior. However, as Hancock (2009) argued, the combination of these moments, in measures such as coefficient of variation (COV), tend to be even more helpful. The latter author suggested that COV was an especially useful diagnostic in terms of predicting progress toward incipient failure. To relate this element to a dimension previously discussed, constant COV output would be one effective way of circumventing surprise in systems that are experiencing slow but consistent perturbation. The next logical step, in assessing characteristics such as brittleness, adaptability, and resilience, is to move from the statics of statistics to the display of the unfolding system dynamics, as epitomized in the body of research on non-linear dynamics. It is one of the present recommendations that this form of relevant information extraction and then its dynamic display in the flight deck can help ameliorate some of the more adverse aspects of unexpected and unanticipated events. This is most especially the case when such events are relatively slow in developing; that degree of temporal development being matched against the capacity of an experienced flight-crew to respond.

13.15 Unexpected Events

When considering the potential unexpected events there is a long-standing problem with terminology being inconsistent between different sources and investigations. The definition that we have adopted here is that an unexpected event is: "*(a) An event incongruent with expectations as determined by base rate probabilities (average probability of event occurring) and the contextual information available; may be normal, abnormal, or emergency in nature; it may also be frequent, infrequent, or novel, or (b) the absence of an expected event."* A taxonomy of the terms used in various studies of unexpected events is presented here in Figure 7 and Table 1. These observations are taken from the work of Kochan et al. (2005). Events here are divided in terms of severity, frequency, and the degree of unexpectancy involved. Events may be divided, as a first pass set of differentiations into normal, abnormal, or emergency. Similarly, we can then sub-divide them into categories that can be labelled common, unusual, or entirely novel; and, they may be expected or unexpected. This provides the tree structure as shown in Figure 7 below. Clearly, these forms of differentiation also provide an intrinsic hierarchy of threat and risk. Novel, emergency, and unexpected conditions are those which the present overall Project is designed to address. However, we explicitly recognize here that such a combination represents only one of many multiple possible combinations, which themselves represent varying degrees of threat to operational safety and efficiency. What remains to be understood is whether such forms of descriptive taxonomy can be elaborated into more thorough descriptions and lead to insights as to how one state then potentially leads to another. It is toward this understanding that other tasks within the present project, and most especially that inquiry elicited from subject matter experts, is directed.

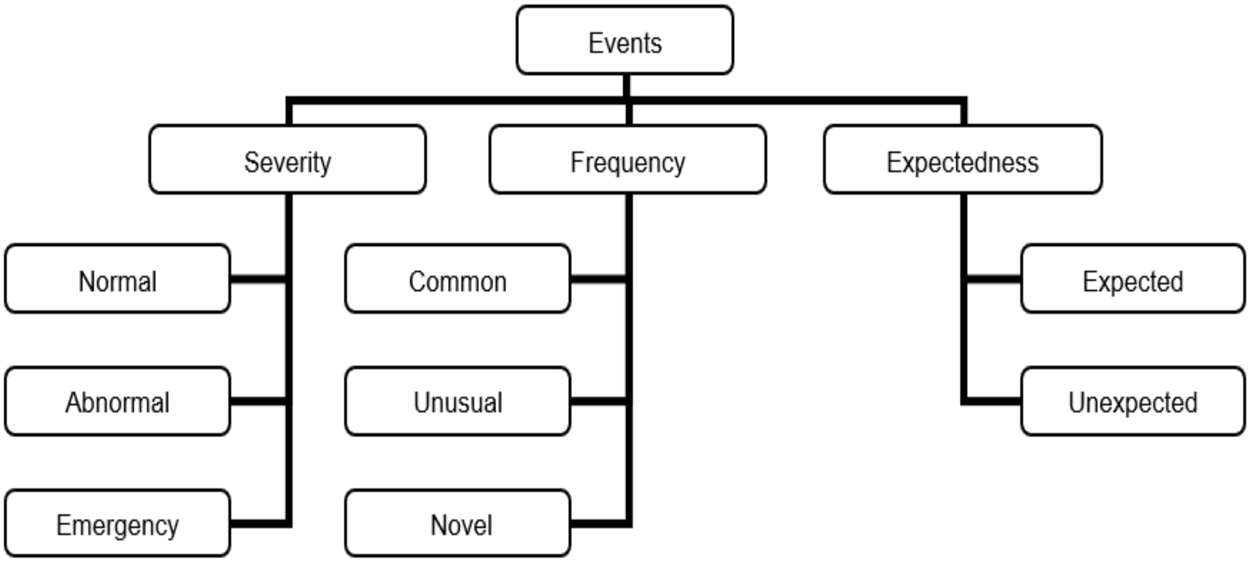

Figure 6. A Descriptive Taxonomy of Events in Terms of Severity, Frequency, and Expectedness.

Table 1. Taxonomy of Events, Adapted from Kochan, Breiter, and Jentsch (2005)

Event Feature (Synonyms)	Working Definition
Severity of Event	
Normal (Nominal or Routine)	The status quo or homeostasis of the system consisting of a series of expected events leading to steady state operating conditions.
Abnormal (Off Nominal or Non-Normal)	An event requiring the use of additional equipment, standby systems, or resulting degraded system (human and/or machine) capabilities.
Emergency Event	A potentially hazardous event requiring timely resolution of the state; may evolve into a normal or abnormal state.
Frequency of Event	
Common Event (Frequent)	Events occurring with a high base rate; most often normal in nature, although may be abnormal or emergency events.
Uncommon Event (Unusual, Rare, Non-Frequent, or Non-Routine)	A seldom occurring event, varies with base rate of occurrences; may be normal, abnormal, or emergency in nature.

Event Feature (Synonyms)	Working Definition
Novel Event	An event which has not been known to happen in the past; may be normal, abnormal, or emergency in nature.
Expectedness of Event	
Expected Event (Usual or Routine)	An event congruent with expectations of human and machine systems, and environment based on some prior knowledge, information, or preparatory information; may be normal, abnormal, or emergency in nature; it may also be frequent, infrequent, or novel.
Unexpected Event	(a) An event which occurs without the affordance of any prior information, cues, or warnings that could function to change (increase or decrease) the probability weighting of the event; may be normal, abnormal, or emergency in nature; it may also be frequent, infrequent, or novel (b) the absence of an expected event.

While these types of events are often the precursors to negative outcomes, it is possible for pilots and other crew to respond in such a way that a negative outcome is averted. In fact, through resilience, the possible adverse influences of such event may be mitigated or avoided altogether.

13.16 Summary of Working Definitions

This section summarizes the current definitions of the primary points of interest throughout this paper (i.e., *resilience, adaptability, cognitive flexibility, decision-making, recognition, novelty, startle, surprise, stress, safety, monitoring, robustness, brittleness, and unexpected events*). Please note that although these definitions are based on the current literature, they are continuously evolving with the goal to ensure a well-operationalized variable for adequate measurement.

*Adaptability***.** We define adaptability as, "being able to strategically return to an original state after an unexpected event" after Hancock & Cruit (2020).

*Brittleness***.** Woods (2015) definition of brittleness has been adopted as "a rapid fall off or collapse of performance that occurs when events push a system beyond its boundaries for handling changing disturbances and variations."

*Cognitive Flexibility***.** According to Spiro (1995), "cognitive flexibility is the ability to spontaneously restructure one's knowledge in many ways, in adaptive response to radically changing situational demands" and has been adopted as our working definition.

*Decision-Making***.** We consider decision-making as a systematic approach to the mental process used to consistently determine the best course of action in response to a given set of circumstances.

Metacognition. The ability to monitor one's current level of understanding and decide when it is and when it is not adequate. In other words, it is the awareness of one's knowledge and is a skill which can be used to control and manipulate cognitive processes.

Monitoring. The working definition for monitoring is adopted from the FAA's Crew Resource Management Advisory Circular (AC 120-51E) is "the observation of the aircraft's flight path and systems and actively cross-checking the actions of other crewmembers."

Novelty. Novelty is considered to be "a property of a stimulus that has not been previously presented to or observed by and is thus unfamiliar to the subject" per Gordon & Luo (2011).
Recognition. The working definition for recognition in the context of this project is, identifying something totally with sense, perception, awareness and/or behavior.

Resilience. Although this paper does not address resilience as our primary focus nor a means to an end, we begin with resilience as it encompasses other terms. Thus, our definition of resilience with a specific focus on the system is the systemic capacity to change as a result of circumstances that push the system beyond the boundaries of its competence envelope. The system may have to amend some, or even all of its goals, procedures, resources, roles, or responsibilities. As a result of those changes, the work system then expresses a revised competence envelope. In effect, it becomes a different system. In addition, we also consider a definition of resilience that pertains directly to the human in the system: "the ability to adapt to changing circumstances by attaining a differing form of operational stability through situation assessment, self-review, decision making, and action." By focusing on the human in the system, we can consider human behavior and implications for training these qualities of resilience as outlined in the definitions.

Robustness. We will use the definition from Hoffman & Hancock (2017) for robustness as, "the ability of the system to operate within its normal operating boundaries when it is perturbed."

Safety. The working definition of safety is from ICAO (2018); "the state in which risks associated with aviation activities, related to, or in direct support of the operation of aircraft, are reduced and controlled to an acceptable level" (p. vii).

Startle. For the purposes of this study, we define startle as, "a physiological reflex reaction to a sudden, intense stimulus triggering an involuntary physiological response to include eye blink, increased heart rate, and increased tension of the muscles."

Stress. "Stress is a response to threatening situations that involves biological, cognitive, behavioural, and emotional components" is adopted as our definition of stress from Dismukes et al. (2015).

Surprise. Surprise is defined as, "an emotional and cognitive response to unexpected and difficult to explain events" as described by Landman et al. (2017).

Unexpected Event. An unexpected event is (a) An event incongruent with expectations as determined by base rate probabilities (average probability of event occurring) and the contextual information available; may be normal, abnormal, or emergency in nature; it may also be frequent, infrequent, or novel, or (b) the absence of an expected event (Kochan et al., 2005).

At present, we can draw certain linkages between the various constructs we have surveyed. For example, brittleness and adaptability are essential antitheses. The more brittle a system is, the less adaptable, and vice-versa. Resilience is a property that exerts its effects at the boundaries of adaptability. High resilience denotes the capacity to readily establish a new normal, low resilience its opposite. When any such 'new normal' has been established, adaptability and brittleness apply in the same way to the new normal as they did to the prior or previous operational state. Hysteresis we take to be an index of adaptability. High levels of hysteresis connote highly adaptable systems and again, vice-versa. Startle, surprise, novelty, unexpected events are all temporal characteristics of the mission space. They overlap, but only to a certain degree. Surprise tends to covary with unexpected events but not completely. Startling stimuli are also often surprising, but not always necessarily so; particularly after the recognition of the startle inducing event has taken place. Novelty is quite often link to unexpected events, but again not in a fully deterministic manner. Cognitive capabilities, such as monitoring, recognition, decision-making, and cognitive flexibility denote the responses that can be taken by the alert and trained crew, or other members of the airspace system. Factors such as stress, fatigue, and excessive cognitive overload or underload, are each facet which interfere with this capacity to respond effectively. Safety represents the state of ambient risk, and by inversion the momentary health of the system in respect of the challenges it faces. Through these emerging and linked concepts, we are able to create differing theories and models which subsume effective response to unanticipated events (and see Landman et al. 2017). These new and exciting approaches are now being promulgated and are certain to have important impacts on training at all levels of the airspace system, We this conclude here that resilience training is a proactive approach through which to prevent incidents, as opposed to the more common reactive methods which prevent incidents from being repeated but not from occurring in the first place. The elements of our Project which follow are predicated upon this critical assumption.

14. Conclusions based on the Findings

The goal of the present sequence of research inquiries is to present to the FAA recommendations and guidance that the FAA can reference in developing resiliency training for novel, unexpected, ill-defined, and otherwise surprising events. This not only relates to problems in aircraft operations including malfunctions, but also teamwork breakdowns and ATC and communication lapses which can also drive events. This section discusses an example of a recent research effort, the present gaps within the literature, and then showcases several examples of flightcrew behaviors from positive outcome events that are cross-referenced with Safety II constructs from this assessment.

14.1 A Current Research Study Example

As a follow-on study to the Pruchnicki et al. (2019) effort, researchers performed a simulator-based experiment to examine aspects of resilient behaviors. The purpose of this task was to observe professional flightcrews undergoing simulator training/evaluation to determine the frequency of resilient behaviors as described in Pruchnicki, et al. (2019). This study task was performed by observing flight training and checking conducted for a jet fractional type operation under 14 CFR Part 91 Subpart K. The simulator observations were part of the operator’s yearly required training and had the format of Line Orientated Flight Training (LOFT). That is, the flightcrews were given a scenario that offered challenges to manage from both a technical and Crew Resource Management (CRM) perspective. This was believed to represent new opportunities to potentially demonstrate resilient behaviors.

The scenarios were designed by the operator with the goal to offer various experiences that would challenge the flightcrews in different ways, but all provided an opportunity to demonstrate resilient behaviors. The foundational basis for the development of the individual storylines came from numerous sources of data. For example, a few of which were based on company safety reports, incident reports and NTSB reports. The goal was to provide a realistic scenario experience that was challenging while providing one that offered a valuable experience to enhance learning and avoid negative training. The simulators that were used in this training/checking event were operated by a contracted professional flight-training center and were full motion FAA certified level D1 simulators. These scenarios were intended to provide the flightcrews with unexpected events that they could react to challenge their technical and CRM skills which may employ resilient behaviors.

This task consisted of two distinct parts. First, researchers observed flightcrews in real time from the rear seats in the simulator for possible indications of observed resilient behaviors. These flights were part of a FAA approved training program as part of recurrent training for this operator and the aircraft were business jets requiring two pilots using level D full motion flight simulators. Researchers used a data collection sheet in order to make data collection in real time more orderly and successful. Although at times, the data collection in real time was challenging, as these are dynamic events, the simulations were occasionally stopped for teaching moments, which facilitated the opportunity to capture more extensive notes. These training events were four hours in length with a break after two hours. The break also afforded a chance to capture more data, thereby completing our observations more extensively. These scenarios did not encompass the entire four-hour training period, as other maneuvers were required to meet the standards set forth by the FAA and were not included in our observational notes. Researchers used this additional time as well to complete notes. Second, researchers analyzed the data to determine which characteristics of resilience were seen the most and how this might have related to the crew being resilient.

14.1.2 Limitations of this Study:

The opportunity to conduct this study allowed for valuable insight into pilots' reactions to unexpected events. Due to the matter that the data were collected during real-world training events, the following limitations were necessary:

- The most overarching study limitation was that researchers were not allowed to interact with the crew including post simulator session interviews. Because of the requested limitations with the degree of intrusiveness, researchers were not allowed to interview the flightcrews once they were finished, as this time is used for the required crew debriefing. Essentially this also prevented the researchers from capturing demographic information about the flightcrews, making it impossible to analyze these variables that may relate to resilience.

- The inability to record the simulator session for later review.

- Researchers were not allowed to design the LOFT scenarios.

- Flightcrews were allowed to choose which scenario they would like to perform during the LOFT. However, specifics such as the exact nature and timing of the unexpected event was unknown to the crew during their simulator time.

In many cases, pilots make hundreds of decisions per flight but when encountering an abnormal or emergency situation, the occurrence can increase significantly. Although many unexpected events have procedures and/or experienced training to help flightcrews mange them successfully. Unfortunately, there are those that do not. For example, in some cases, prescribed procedures need a degree of modification because either they are not fully applicable, or time is too short to complete the entire procedure (For example the US Airways' flight 1549 landing on the Hudson River). During these times,

---

1 Level D simulators are those that have 6-axes of motion and sophisticated visuals with at least 150 degrees of viewing.

pilots are expected to evaluate the situation and respond quickly. Unfortunately, depending on the flightcrew, Pruchnicki and colleagues observed that some of these events had better outcomes than others since some resulted in an accident event during the LOFT. Thus, the question is - why are some flightcrews better able to respond to unexpected events than others? One explanation could be that some flightcrews are simply better at their tasks than other pilots, either due to greater experience or natural ability. This can factor in when analyzing performance; however, note that unexpected events sometimes result in a negative outcome even if the flightcrew is highly experienced (e.g., a loss of control after a flight control malfunction). Situation-specific factors, such as time pressure and the type of unexpected event itself, may also influence performance. Thus, experience/ability and the situation-specific factors likely interact to influence the outcome. The goal of this research is to empower pilots (regardless of experience or innate ability) to respond in an adaptive way to all types of unexpected event situations. As discussed, one promising avenue or pathway is the construct of resilience engineering.

Resilience Engineering (RE), through understanding and promoting resilient skills and behaviors that can be applied across a wide range of unexpected events and flight crew ability levels. Resilience as a concept is about enhancing people's adaptive capacity so that they can counter unexpected threats beyond their prepared abilities (Dekker & Lundstrom, 2006). This is an important distinction. Pruchnicki and colleagues believe that resilience is a set of skills that is utilized when events occur that go beyond normal emergency/incident training. Therefore, a resilient skill set should be one that is highly adaptable and can be applied to challenging events.

To capture resilience in action, Pruchnicki and colleagues observed 17 flightcrews completing simulator training containing some element of a novel or unexpected event. The frequency of times they used each of six hallmark resilient behaviors was tallied (208 examples). Depending on the crew and the scenario chosen, some of these resilient behaviors were seen more than others. The results of these observations indicated that *monitoring* and *responding* occurred much more frequently than the other behaviors. This is not surprising since many times in both normal and non-normal flight, these are frequently required to manage most situations. The very nature of aviation and pilot training lends itself to the use of these two strategies. Pilots should monitor the aircraft status in order to fully understand what is happening before responding. Once the situation has been fully understood and the aircraft is under control, the crew can then follow memory items and guidance from the QRH2 if applicable. As mentioned above, learning was less commonly seen due to the researchers' inability to interview the flightcrews after the simulator session so that we could discover what they brought from their past to the current training event. The results indicated that *responding* was used with significantly higher frequency than all other variables except monitoring, $p < .002$. Moreover, as previously mentioned, learning was observed significantly less frequently than all the other skills except anticipating, $p < .001$. There were no other statistically significant differences in frequency of the behaviors observed. Despite the meaningful statistical findings, the tallying of just the cornerstones is not enough to fully understand resilient behaviors to the level or granularity needed to determine more discerning and specific behaviors. These include the skills, methods, knowledge, strategies utilized to manage these situations. These behaviors, once better understood, will offer researchers a deeper understanding into resilience. Pruchnicki and colleagues suggest that future research should include classroom training on resilient behaviors and simulator testing against a control group to assess outcomes. This will enable researchers

---

2 QRH – a manual that has all of the checklist for mechanical problems with the aircraft.

to better understand whether resilient behaviors can be trained and might suggest what training methods should be utilized.

14.2 Gaps within the Literature

While Pruchnicki and colleagues (2019) covered much of what is understood in the way of resilience, one thing they did not explore was team resilience. Training interventions, aimed at improving an individual's ability, will certainly be useful. However, team interaction is also one of the major factors in both causing and preventing unwanted outcomes. This form of evaluation becomes especially complicated as autonomy increases, and the team members may not all be human operators but also include many technical support systems. A resilient pilot, using a resilient system, may be operating under a different mental model and thus their combined effort may not prove to be an additive advantage. Indeed, collaboration may be worse, than the efforts of each one alone. For this reason, team resiliency is an issue that we look to address in the present work.

The current paragraph addresses some issues with measuring individual and systems resilience but offers insight into the literature on how one might measure characteristics of resilience. Pruchnicki and his colleagues (2019) addressed the fact that it could be difficult to measure whether or not any training interventions, meant to increase resilience, actually prove to be effective. After all, resilience represents a complicated trait which cannot be measured directly. However, there are some indirect methods to determining the characteristics of resilience. Hoffman and Hancock (2017) have suggested using time as a central element of measuring resilience. For instance, they have pointed out that the time it takes a system to notice a problem, come up with a solution, and enact that solution, can very well be an indication of the relative ease with which the system responded to a threat. As for the human operator, behavioral skills could be measured as proxies for resilience. These include confidence, flexibility, and adaptability. Furthermore, Pruchnicki and colleagues (2019) recommended determining which crew behaviors were evidently related to successful outcomes and using those behaviors as a measure of resilience. The relative efficacy of training interventions in those areas could then be tested to determine whether resilience engineering is a successful and effective enterprise. The following paragraphs discuss a follow-on study to Pruchnicki and colleagues (2019) work.

14.3 Good Outcome Events Cross-Referenced with Research Findings

In the rare events where the unexpected does occur, machines rely essentially on their programming, whereas humans possess flexibility to respond to an entirely new circumstance. And, as we previously observed, modern adverse events in aviation often are almost entirely novel. Therefore, the human quality of resilience, or the ability to respond to interruptions outside of what has been pre-prepared for, is one of the reasons why human pilots will be needed, at least for the foreseeable future. As discussed earlier, this is in-line with the thought process of the Safety II literature (Schwarz et al., 2016).

By identifying the positive behaviors exhibited by aircrew that led to positive outcomes, this may point to patterns of repeatability, trainability, and ultimately maintaining safer aviation systems. Consider the following three events in relatively recent history (see Table 2), which provide examples of how pilots and crew exercised appropriate responses when faced with the unexpected. A team of three human factors experts independently analyzed the four events displayed in Table 2 and then coded all positive behaviors revealed by aviation personnel and passengers associated with the events. Then, the positive behaviors were categorized by Safety II constructs from this research assessment (i.e., decision-making strategies, metacognition, cognitive flexibility, and expertise). Within these examples, it is evident that crewmembers relied not only on standard flight training to critically maneuver through these events, but CRM training and lessons learned through previous live flights as well. Take for example, US Airways Flight 1549. There is an exhaustive list of positive behaviors associated with the captain, crew, and even passenger's decision making. As evident in the National Transportation and Safety Board's investigations of these incidents (NTSB, 1989; NTSB, 1989; NTSB, 2010; NTSB, 2016), there is more documentation of the positive actions associated with crew's decision making and cognitive flexibility in the more recent events, such as US Airlines Flight 1549, whereas earlier events neglected to describe the positive actions pilots and flightcrew demonstrated. This observation demonstrates the need for a continuation and a stronger standardization of documenting positive actions in all aviation events

Table 2. Positive Behaviors Associated with Aviation Events

Event	Date	Safety II Constructs from Research Assessment	Positive Behaviors
United Airlines Flight 232 (Sioux Gateway Airport)	July 19, 1989	Cognitive Flexibility	The ability for aircrew to quickly learn to work with new teams such as the National Guard.
		Decision Making Strategies	Positive communication among the aircrew and the respect of each crewmember.
		Expertise	High level of domain and judgment expertise among the crew.
Aloha Airlines Flight 243	April 28, 1988	Expertise	High level of flight expertise between the captain and first officer.
		Cognitive Flexibility	The ability to flex communication styles from verbal to nonverbal due to ambient noise
		Decision Making Strategies	Thinking quickly under time pressure and then responding quickly.
US Airways Flight 1549	January 15, 2009	Decision Making Strategies	Effective communication and coordination between captain and first officer.
		Decision Making Strategies	The allocation of tasks to passengers and their openness to perform tasks.
		Cognitive Flexibility	Flight attendants were able to flex the evacuation commands by instructing those who were able to jump over seats.
		Cognitive Flexibility	The ability to mentally adjust checklists in order to perform under time pressure, applying successes from past experiences.
		Metacognition	Captain being aware that he had the confidence to land the plane in the Hudson River.
Southwest Flight 3472	August 27, 2016	Decision Making Strategies	Decision by the crew to complete the SWA Engine Fire or Engine Severe Damage or Separation Checklist; Decision by Captain to delegate tasks to First Officer and cabin crew members.
		Cognitive Flexibility	Crew coordination and the ability to change the mode of communication throughout the event (e.g., crew switched from verbal to non-verbal communication when the noise level made words inaudible.)

15. Summary

Traditional training has sought to provide crewmembers the experience in practicing known procedures to known events. That is, those events that deemed most likely or at least historically so. In this context, normal training as performed today can be very useful in allowing for this practice in the safe environment of a simulator. However, many of the types of system failures seen today occur where no specific procedure is known. These events are unexpected and may surprise the human. One common characteristic of these types of occurrences is that they cannot specifically be trained for and are outside the scope of normally trained and practiced situations. We typically train for events in which there are procedures; therefore, how do we train for the element of surprise?

The expansive literature on expertise provides insight into how pilots may think and behave when faced with a problem (e.g., an unexpected event). Developing expertise at managing an unexpected event is a product of more than actual flight hours, simulator time, and classroom training and we know from the literature that being an expert in one domain does not transfer to other domains. The pilot with 30,000 flight hours who has only flown on the Eastern Coast of the United States may not recognize the subtle cues of a brewing sandstorm such as the less experienced pilot who has clocked only 300 hours in the deserts of Arizona. But as previously mentioned, it is not possible to train pilots to become experts at unknown procedures. However, it may be possible to train pilots to develop expertise in adaptability, cognitive flexibility, and resilience.

With the need for training pilots in cognitive flexibility, adaptation and resilience comes the need for effective training measures. It has been suggested that resilience is a set of skills that is hard to measure and some feel that it is very difficult to train. However, Hollnagel (2009) explains that the four cornerstones of resilience (e.g., anticipation, monitoring, responding and learning) may represent areas of focus for future training programs.

Not only is training for resilient behaviors challenging, but the design of studies to examine their applicability is difficult as there is no agreed upon method of measuring resilience and thus a metric to establish which methods may be better suited than others. One possible method of measuring resilience is based on time to design and implement novel solutions to unexpected events. This time-based approach considers the reaction time of experts versus novices. Cognitive explanation includes how short-term memory and long-term memory are utilized and accessed when faced with these types of events. Experts have more conceptual chunks available to both long and short-term memory, more cues defining each chunk, more interrelationships among the cues and chunks, and more pathways available to access all, or parts of the chunks as needed, to solve a problem or make a decision. One possibility for training resilience is to capitalize on the breadth of experience experts have in relation to novices. That is, to train such that it allows pilots to critically think through novel situations by linking them to past experiences. These scenarios can start at the novice level and extend to experienced pilots. With a greater experience level, this is more practical for experienced pilots versus novices who have less experience to draw from. Although individual expertise and resilience is an important consideration, the characteristics of team resilience is an underappreciated area of research. It is believed that communication solutions such as those taught in joint training exercises, specifically crew resource management (CRM) as being helpful in training crews to be more effective in demonstrating resilient behaviors. One area of training that could be considered an extension of CRM is threat and error management (TEM) as its current iteration is centered on recognizing that errors will occur and the focus should be on error tolerance which first involves error recognition.

The implications of the findings from this report suggest the need for collecting more qualitative data with both expert and novice pilots to gain a better understanding of the differences in how experts and novices make decisions during unexpected events. In addition, more research is needed in assessing whether pilots can be trained to possess the skills and behaviors necessary to recover from an unexpected event. These studies would involve first developing experimental studies to test our theoretical measures, and then developing and testing a training program aimed at increasing pilot performance during unexpected events. Ultimately, work in this area would provide pilots and crew with the knowledge and behaviors to maintain safer air operating systems under known and unknown circumstances.

The results of this effort conclude the following findings and recommendations for future work:

- Key findings and possible ways forward:
    - There is no empirically validated training program specifically aimed at increasing pilot resilience during unexpected events.
    - There is no empirically validated way to specifically assess pilot resilience during unexpected events.
    - We believe looking at judgment expertise is a possible way forward. Judgement expertise has been validated and assessed.
    - Examining the relationship between experts and novices may lead to developing resilience training and assessment.
    - Training should focus on critical thinking skills such as leveraging past experiences.
    - Case studies showed that successful pilots demonstrated the following behaviors (referenced in Figure 3 and associated texts):
        - Expertise
        - Critical thinking
        - Metacognition
        - High self-efficacy
        - Cognitive flexibility
        - Leveraging past experiences
    - Self-Efficacy may be a predictor of resilience during unexpected events.
        - Assessing self-efficacy during a scenario script may guide training programs on raising pilots' confidence in their own expertise.
    - Scenario-based training and scenario-based assessment has shown to increase pilot expertise. This is a promising way forward.
    - Time-based measurements of resilience paired with Hollnagel's four cornerstones of resilience may be a possible way forward.
        - But, researchers need to study the knowledge, skills, methods, and strategies needed to successfully manage unexpected events. Just focusing on the 4 cornerstones of resilience is not enough.

- How these findings will contribute to our future tasks:
    - Task 3 will focus on understanding more than just the 4 cornerstones of resilience. Through our knowledge elicitation interviews, we will focus on the skills that Shawn recommends such as pilots' knowledge, skills, decision making, critical thinking, and other strategies used to manage unexpected events.
    - Phase 2 will focus on designing scenario-based training, and future research should test that training against a control group to determine if resilience training does in fact increase pilot performance in a simulated flight.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74. https://doi.org/10.1037/0033-2909.101.1.41

Ali, S., Maciejewski, A. A., Siegel, H. J., & Jong-Kook Kim. (2004). Measuring the robustness of a resource allocation. *IEEE Transactions on Parallel and Distributed Systems*, *15*(7), 630–641. https://doi.org/10.1109/TPDS.2004.24

Alliger, G. M., Cerasoli, C. P., Tannenbaum, S. I., & Vessey, W. B. (2015). Team resilience: How teams flourish under pressure. *Organizational Dynamics, 44*(3), 176–184. https://doi.org/10.1016/j.orgdyn.2015.05.003

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. https://doi.org/10.1016/S0022-5371(83)90201-3

Barnett, C. K. and Pratt, M. G. (2000). From threat-rigidity to flexibility - Toward a learning model of autogenic crisis in organizations. *Journal of Organizational Change Management*, *13*(1), 74–88. https://doi.org/10.1108/09534810010310258

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology*, *39*(1), 15–22. https://doi.org/10.1080/00221309.1948.9918159

Bhamra, R., Dani, S., & Burnard, K. (2011). Resilience: The concept, a literature review and future directions. *International Journal of Production Research*, *49*(18), 5375–5393. https://doi.org/10.1080/00207543.2011.563826

Broadbent, D. E., & Ladefoged, P. (1959). Auditory perception of temporal order. *The Journal of the Acoustical Society of America*, *31*(11), 1539-1539. https://doi.org/10.1121/1.1907662

Bureau Enquêtes-Accidents (1998). *Rapport relatif à l'abordage survenu le 30 juillet 1998 en baie de Quiberon entre le Beech 1900D immatriculé F-GSJM exploité par Proteus Airlines et le Cessna 177 immatriculé F-GAJE* [Report of the collision occurring on July 30, 1998 in Quiberon Bay between Beechcraft 1900D registered F-GSJM operated by Proteus Airlines and the Cessna 177 registered F-GAJE]. https://www.bea.aero/docspa/1998/f-je980730/htm/f-je980730.htm

Callister, W. D., & Rethwisch, D. G. (2007). *Materials Science and Engineering: An Introduction* (10th edition). John Wiley and Sons.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. F., (1995). Defining competencies and establishing team training requirements. In R. Guzzo & Salas, E. (Eds.), *Team effectiveness and decision making in organizations.* (pp. 333–380). Jossey Bass.

Cañas, J. J., Antolí, A., Fajardo, I., & Salmerón, L. (2005). Cognitive inflexibility and the development and use of strategies for solving complex dynamic problems: effects of different types of training. *Theoretical Issues in Ergonomics Science, 6*(1), 95–108. https://doi.org/10.1080/14639220512331311599

Chandra, D. C., & Markunas, R. (2016). Line pilot perspectives on complexity of terminal instrument flight procedures. *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, 1–10. https://doi.org/10.1109/DASC.2016.7778012

Chandra, D. C., Lennertz, T., & Cardosi, K. (2019). Flight deck human factors issues in vertical and lateral deviations during north atlantic flight operations. *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*, 1–9. https://doi.org/10.1109/DASC43569.2019.9081742

Chandra, D., Sparko, A., Kendra, A., & Kochan, J. (2020). Operational complexity in performance-based navigation arrival and approach flight operations. In N. Stanton (Ed.), *Advances in human aspects of transportation* (pp. 531–538). Springer International Publishing. https://doi.org/10.1007/978-3-030-50943-9_67

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

Chi, M. T., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Lawrence Erlbaum Associates, Inc.

Cohen, I. (2008). Improving time-critical decision making in life-threatening situations: Observations and insights. *Decision Analysis, 5*(2), 100–110. https://doi.org/10.1287/deca.1080.0111

Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, *18*(2), 76–82. https://doi.org/10.1002/da.10113

Cruit. J. (2016). *Predicting general aviation pilots' weather-related performance through a scenario-based written assessment.* [Doctoral Dissertation, Embry-Riddle Aeronautical University]. https://commons.erau.edu/edt/198

Cruit, J., Clapp, S., Kochan, J., & Hancock, P.A. (2019). Perspectives toward training resiliency during unexpected events in air carrier operations. *Presented at the 20th International Symposium of Aviation Psychology*.

Dalziell, E. P., & McManus, S. T. (2004). Resilience, vulnerability, and adaptive capacity: Implications for system performance. *International Forum for Engineering Decision Making* (IFED).

Davis, M. (1984). The mammalian startle response. In R. C. Eaton (Ed.), *Neural mechanisms of startle behavior* (pp. 287–351). Springer US. https://doi.org/10.1007/978-1-4899-2286-1_10

Dekker, S. W. & Lundstrom, J. (2006). From threat and error management (TEM) to resilience. *Human Factors and Aerospace Safety, 6*(3), 261–274. https://trid.trb.org/view/842420

Dekker, S., & Pruchnicki, S. (2014). Drifting into failure: Theorising the dynamics of disaster incubation. *Theoretical Issues in Ergonomics Science*, *15*(6), 534–544. https://doi.org/10.1080/1463922X.2013.856495

Deming, W. E. (1953) Management's responsibility for the use of statistical techniques in industry. *Global Journal of Advanced Research*. http://gjar.org/articles/DEMING-S--MANAGEMENT-S-RESPONSIBILITY-FOR-THE-USE-OF-STATISTICAL-TECHNIQUES-IN-INDUSTRY--REVISITED--

Dennis, J. P., & Vander Wal, J. S. (2010). The cognitive flexibility inventory: Instrument development and estimates of reliability and validity. *Cognitive Therapy and Research*, *34*(3), 241–253. https://doi.org/10.1007/s10608-009-9276-4

Dismukes, R. K., Kochan, J. A., & Goldsmith, T. E. (2018). Flightcrew errors in challenging and stressful situations. *Aviation Psychology and Applied Human Factors*, *8*(1), 35–46. https://doi.org/10.1027/2192-0923/a000129

Dismukes, R. K., Goldsmith, T. E., & Kochan, J. A. (2015*). Effects of acute stress on aircrew performance: Literature review and analysis of operational aspects* (NASA/TM-2015-218930). NASA Ames Research Center. https://human-factors.arc.nasa.gov/flightcognition/Publications/NASA_TM_2015_218930-2.pdf

Dismukes, K., Young, G., & Sumwalt, S. (1998). *Cockpit interruptions and distractions: Effective management requires a careful balancing act* (Document ID 2002006298). NASA Ames Research Center. https://ntrs.nasa.gov/citations/20020062698

Dumitrescu, A.-T., Oltean, E., Merezeanu, D., & Dobrescu, R. (2017). Emergence in hierarchical complex systems structured as social networks. *2017 21st International Conference on Control Systems and Computer Science (CSCS)*, 426–431. https://doi.org/10.1109/CSCS.2017.66

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*(3), 183–201. https://doi.org/10.1037/h0047707

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*(1), 273–305. https://doi.org/10.1146/annurev.psych.47.1.273

Ericsson, K. A., & Smith, J. (Eds.). (1991). *Toward a general theory of expertise: Prospects and limits*. Cambridge University Press.

European Aviation Safety Agency. (2016). *Terms of reference for a rulemaking task: Evidence-based and competency-based training* (RMT.0599, TOR Issue 1). https://www.easa.europa.eu/document-library/terms-of-reference-and-group-compositions/tor-concept-paper-rmt0599

European Aviation Safety Agency (2017a). *Acceptable means of compliance and guidance material to Anex III-Part-ORO*. https://www.easa.europa.eu/document-library/acceptable-means-of-compliance-and-guidance-materials

European Aviation Safety Agency (2017b) *Crew Resource Management in Practice*.

European Aviation Safety Agency (2020). Easy access rules for flightcrew licensing (Part-FCL). https://www.easa.europa.eu/easy-access-rules-flight-crew-licencing-part-fcl

Directorate-General for Research and Innovation (European Commission). (2018). *Aviation safety: Challenges and ways forward for a safe future*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/417337

Farjadian, A. B., Thomsen, B., Annaswamy, A. M., & Woods, D. D. (2021). Resilient Flight Control: An Architecture for Human Supervision of Automation. *IEEE Transactions on Control Systems Technology*, *29*(1), 29–42. https://doi.org/10.1109/TCST.2019.2959542

Federal Aviation Administration. (2015). *Stall prevention and recovery training* (Advisory Circular 120-109A). https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/1028646

Federal Aviation Administration. (2000). *FAA system safety handbook: Principles of system safety.* https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/risk_management/ss_handbook/media/chap3_1200.pdf

Federal Aviation Administration. (2004). *Crew resource management training* (Advisory Circular 120-51E). https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentid/22879

Federal Aviation Administration. (2009). *Pilot's handbook of aeronautical knowledge* (FAA Handbook 8083-25B). Skyhorse Publishing Inc. https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/phak/

Federal Aviation Administration. (2009). *Risk management handbook* (FAA Handbook 8083-2). https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/media/faa-h-8083-2.pdf

Federal Aviation Administration. (2017). *Standard operating procedures and pilot monitoring duties for flight deck crewmembers* (Advisory Circular Nr. 120-71B) https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentid/1030486

Faerevaag, C. L., Rivera, J. A., Jimenez, C. A., & Jentsch, F. (2018). *A review of current CRM training and evaluation techniques implemented by U.S. Air Carriers* (FAA-CRA 17G-001).

Falkland, E. C., & Wiggins, M. W. (2019). Cross-task cue utilisation and situational awareness in simulated air traffic control. *Applied Ergonomics*, *74*, 24–30. https://doi.org/10.1016/j.apergo.2018.07.015

Field, J. N., Boland, E. J., van Rooj, J. M., Mohrmann, J. F. W., Smeltink, J. W. (2015). *Startle effect management* (NLR-CR-2018-242). European Aviation Safety Agency. https://www.easa.europa.eu/document-library/research-reports/easarepresea20153

Frazzini, R. (2001). Technology impact: Some thoughts on deskilling and design responsibility. *IEEE Control Systems Magazine*, *21*(1), 8–12. https://doi.org/10.1109/37.898787

Gallopín, G. C. (2006). Linkages between vulnerability, resilience, and adaptive capacity. *Global Environmental Change*, *16*(3), 293–303. https://doi.org/10.1016/j.gloenvcha.2006.02.004

Gibson, C. B., & Birkinshaw, J. (2004). The antecedents, consequences, and mediating role of organizational ambidexterity. *Academy of Management Journal, 47*(2), 209–226. https://doi.org/10.2307/20159573

Gillespie, B. M., Chaboyer, W., Wallis, M., & Grimbeek, P. (2007). Resilience in the operating room: Developing and testing of a resilience model. *Journal of Advanced Nursing*, *59*(4), 427–438. https://doi.org/10.1111/j.1365-2648.2007.04340.x

Glass, B. D., Maddox, W. T., & Love, B. C. (2013). Real-time strategy game training: emergence of a cognitive flexibility trait. *PloS One, 8*(8), e70350. https://doi.org/10.1371/journal.pone.0070350

Gluck, K. A., McNamara, J. M., Brighton, H., Dayan, P., Kareev, Y., Krause, J., Kutzban, R., Selten, R., Stevens, J. R., Voekl, B., & Wimsatt, W. C. (2012). Robustness in a variable environment. In P. Hammerstein & J. R. Stevens, (Eds.), *Evolution and the mechanisms of decision making* (pp. 195–214). Cambridge University Press. https://doi.org/10.7551/mitpress/9556.003.0016

Golden, C. J. (1975). A group version of the Stroop Color and Word Test. *Journal of Personality Assessment, 39*(4), 386–388. https://doi.org/10.1207/s15327752jpa3904_10

Gomes, J. O., Woods, D. D., Carvalho, P. V., Huber, G. J., & Borges, M. R. (2009). Resilience and brittleness in the offshore helicopter transportation system: the identification of constraints and sacrifice decisions in pilots' work. *Reliability Engineering & System Safety, 94*(2), 311–319. https://ideas.repec.org/a/eee/reensy/v94y2009i2p311-319.html

Gordon, C. L., & Luo, S. (2011). The Personal Expansion Questionnaire: Measuring one's tendency to expand through novelty and augmentation. *Personality and Individual Differences, 51*(2), 89–94. https://doi.org/10.1016/j.paid.2011.03.015

Gustafson, L. A., & Orne, M. T. (1963). Effects of heightened motivation on the detection of deception. *Journal of Applied Psychology, 47*(6), 408–411. https://doi.org/10.1037/h0041899

Guastello, S. J. (1995). *Chaos, catastrophe, and human affairs: Applications of nonlinear dynamics to work, organizations, and social evolution.* Lawrence Erlbaum Associates, Inc.

Hamel, G., & Valikangas, L. (2003). The quest for resilience. *Harvard Business Review, 81,* 52–63. https://hbr.org/2003/09/the-quest-for-resilience

Hammond, K R. (2000). *Judgments under stress*. Oxford University Press.

Hancock, G. M., Longo, L., Young, M. S., & Hancock, P. A. (2021). Mental Workload. In *Handbook Of Human Factors And Ergonomics* (pp. 203–226). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119636113.ch7

Hancock, P. A. (1986). The effect of skill on performance under an environmental stressor. *Aviation, Space, and Environmental Medicine*, *57*(1), 59–64. https://pubmed.ncbi.nlm.nih.gov/3510619/

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, *31*(5), 519–537. https://doi.org/10.1177/001872088903100503

Hancock, P. A., Vincenzi, D. A., Wise, J. A., & Mouloua, M. (Eds.). (2008). *Human factors in simulation and training*. CRC Press.

Hancock, P. A. (2009). *Mind, machine and morality: Toward a philosophy of human-technology symbiosis*. Ashgate Publishing, Ltd.

Hancock, P. A., Hancock, G. M., & Warm, J. S. (2009). Individuation: The N = 1 revolution. *Theoretical Issues in Ergonomics Science, 10*(5), 481–488. https://doi.org/10.1080/14639220903106387

Hancock, P.A. (2013). In search of vigilance: The problem of iatrogenically created psychological phenomena. *American Psychologist, 68*(2), 97–109. https://doi.org/10.1037/a0030214

Hancock, P.A. (2014). Automation: How much is too much? *Ergonomics*, *57*(3), 449–454. https://doi.org/10.1080/00140139.2013.816375

Hancock, P.A. (2017). On the nature of vigilance. *Human Factors*, *59*(1), 35–43.

Hancock, P. A. (2021a). Months of monotony—moments of mayhem: Planning for the human role in a transitioning world of work. *Theoretical Issues in Ergonomics Science, 22*(1), 63–82. https://doi.org/10.1080/1463922X.2020.1753260

Hancock, P. A. (2021b). Avoiding adverse autonomous agent actions. *Human–Computer Interaction*, 1–26. https://doi.org/10.1080/07370024.2021.1970556

Hancock, P. A., & Cruit, J. (2020). *The route to resilience.* Manuscript in Submission.

Hancock, P. A., & de Ridder, S. (2003). Behavioral response in accident-likely situations. *Ergonomics*, *46*(12), 1111–1135.

Hancock, P. A., & Dirkin, G. R. (1983). Stressor induced attentional narrowing: Implications for design and operation of person-machine systems. *Proceedings of the Human Factors Association of Canada, 16*, 19–21.

Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2021). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Human Factors*, *63*(7), 1196–1229. https://doi.org/10.1177/0018720820922080

Hancock, P. A., & Volante, W. G. (2020). Quantifying the qualities of language. *PLoS ONE, 15*(5), Article e0232198. https://doi.org/10.1371/journal.pone.0232198

Hassall, M. E., Sanderson, P. M., & Cameron, I. T. (2014). The development and testing of SAfER: A resilience-based human factors method. *Journal of Cognitive Engineering and Decision Making, 8*(2), 162–186. https://doi.org/10.1177/1555343414527287

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Children development and education in Japan* (pp. 262–272). Freeman.

Heichal, G. (1992). *Decision-making in crisis time*. Maarchot IDF Press.

Heinrich H. W. (1931). *Industrial accident prevention: A scientific approach*. McGraw-Hill.

Helmreich, R.L., Klinect, J.R., & Wilhelm, J.A. (1999, May). Models of threat, error, and CRM in flight operations. *Proceedings of the Tenth International Symposium on Aviation Psychology, 10,* 677–682.

Hoffman, R. R., Deal, S. V., Potter, S., & Roth, E. M. (2010). The practitioner's cycles, part 2: Solving envisioned world problems. *IEEE Intelligent Systems*, *25*(3), 6–11. https://doi.org/10.1109/MIS.2010.89

Hoffman, R. R., & Hancock, P. A. (2017). Measuring resilience. *Human Factors, 59*(4), 564–581. https://doi.org/10.1177/0018720816686248

Holbrook, J. B., Stewart, M. J., Smith, B. E., Prinzel, L. J., Matthews, B. L., Avrekh, I., Cardoza, C. T., Ammann, O. C., Adduru, V., & Hull, C. H. (2019). Human performance contributions to safety in commercial aviation. (NASA/TM-2019-220417). NASA Ames Research Center. https://hsi.arc.nasa.gov/publications/NASA-TM-2019-220417-comp.pdf

Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.) (2006). *Resilience engineering: Concepts and precepts, Volume 1*: *Remaining sensitive to the possibility of failure*. Ashgate Publishing, Ltd.

Hollnagel, E. (2009). The four cornerstones of resilience engineering. In C. P. Nemeth, E. Hollnagel, & S. W. A. Dekker (Eds.) *Resilience engineering perspectives, Volume 2: Preparation and restoration* (pp. 117–134). Ashgate Publishing, Ltd.

Hollnagel, E. (2014). *Safety-I and safety-II: the past and future of safety management.* Ashgate Publishing, Ltd.

Hollnagel, E. (2017). *Safety-II in practice. Developing the resilience potentials*. Routledge.

Horstmann, G. (2006). Latency and duration of the action interruption in surprise. *Cognition and Emotion, 20*(2), 242–273. https://doi.org/10.1080/02699930500262878

Ilgen, D. R., & Pulakos, E. D. (Eds.) (1999). *The changing nature of performance: Implications for staffing, motivation, and development.* Jossey-Bass Inc.

Infurna, F. J., & Luthar, S. S. (2016). Resilience to major life stressors is not as common as thought. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 11*(2), 175–194. https://doi.org/10.1177/1745691615621271

International Civil Aviation Authority (ICAO). (2018). Safety management manual *(4th Edition)* (Doc 9859). International Civil Aviation Authority. https://skybrary.aero/sites/default/files/bookshelf/5863.pdf

Jensen, R. S. (1995). *Pilot judgment and crew resource management.* Routledge.

Jensen, R. S., Adrion, J., & Maresh. J. (1987). *Pilot judgment skills test battery*. Report to Clinical Sciences Division, U.S. Air Force School of Aerospace Medicine, Brooks Air Force Base.

Jensen, R. S. and Benel, R. A. (1977). *Judgment evaluation and instruction in civil pilot training*. National Technical Information Service, Final Report (FAA-RD-78-24).

Kahneman, D. (2011) *Thinking, fast and slow.* Farrar, Straus, and Giroux.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*(6), 515–526. https://doi.org/10.1037/a0016755

Kennedy, D. M., Landon, L. B., & Maynard, M. T. (2016). Extending the conversation: Employee resilience at the team level. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 9*(2), 466–475. https://doi.org/10.1017/iop.2016.41

Kinney, L., & O'Hare, D. (2019). Responding to an unexpected in-flight event: Physiological arousal, information processing, and performance. *Human Factors*, *62*(5), 737–750. https://doi.org/10.1177/0018720819854830

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 138–147). Ablex Publishing.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.) (1993*). Decision making in action: Models and methods*. Ablex Publishing. https://doi.org/10.1002/bdm.3960080307

Kleinmann, M. (1993). Are rating dimensions in assessment centers transparent for participants? Consequences for criterion and construct validity. *Journal of Applied Psychology, 78*(6), 988–993. https://doi.org/10.1037/0021-9010.78.6.988

Koch, M. (1999). The neurobiology of startle. *Progress in Neurobiology*, *59*(2), 107–128. https://doi.org/10.1016/s0301-0082(98)00098-7

Kochan, J. (2005). *The role of domain expertise and judgment in dealing with unexpected events.* [Unpublished doctoral dissertation]. University of Central Florida. https://stars.library.ucf.edu/etd/456/

Kochan, J. A., Breiter, E.G., & Jentsch, F. (2005). Surprise and unexpectedness in flying: Factors and features. *Proceedings of the 13th International Symposium on Aviation Psychology*. Wright State University. https://corescholar.libraries.wright.edu/isap_2005/59

Kochan, J. A., Jensen, R. S., Chubb, G. P., & Hunter, D. R. (1997). *A new approach to aeronautical decision-making: The expertise method* (DOT/FAA/AM-97-6). Federal Aviation Administration, Office of Aviation Medicine. https://rosap.ntl.bts.gov/view/dot/40461

Landman, A., Groen, E.L., van Paassen, M. M., Bronkhorst, A. W., & Mulder, M. (2017). Dealing with unexpected events on the flight deck: A conceptual model of startle and surprise. *Human Factors, 59*(8), 1161–1172. https://doi.org/10.1177/0018720817723428

Landman, A., van Oorschot, P., van Paassen, M. M.R., Groen, E.L., Bronkhorst, A. W., & Mulder, M. (2018). Training pilots for unexpected events: A simulator study on the advantage of unpredictable and variable scenarios. *Human Factors, 60*(6), 793–805. https://doi.org/10.1177/0018720818779928

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.

Loukopoulos, L. D., Dismukes, R. K., & Barshi, I. (2001). Cockpit interruptions and distractions: A line observation study. In *Proceedings of the 11th international symposium on aviation psychology* (pp. 1-6). Ohio State University Press. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.4.2145

Luxair Human Factors Training Team. (2013). *Course concept for topic resilience engineering*. CRM-Joint Training.

Magno, C. (2010). The role of metacognitive skills in developing critical thinking. *Metacognition and Learning*, *5*(2), 137–156. https://doi.org/10.1007/s11409-010-9054-4

Martin, W. L. (2019). Crew resource management and individual resilience. In B. G. Kanki, J. Anca, & T. R. Chidester (Eds.), *Crew resource management* (pp. 207–226). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-812995-1.00007-5

Martin, M. M., & Anderson, C. M. (1998). The cognitive flexibility scale: Three validity studies. *Communication Reports, 11*(1), 1–9. https://doi.org/10.1080/08934219809367680

Martin, M. M., & Rubin, R. B. (1995). A new measure of cognitive flexibility. *Psychological Reports, 76*(2), 623–626. https://doi.org/10.2466/pr0.1995.76.2.623

Martin, R. M. (2020). *Electronic structure: basic theory and practical methods*. Cambridge University Press.

Matthews, G., Lin, J., & Wohleber, R. (2017). Personality, stress and resilience: a multifactorial cognitive science perspective. *Psychological Topics, 26*(1), 139–162. https://doi.org/10.31820/pt.26.1.6

Maynard, M. T., Kennedy, D. M., & Sommer, S. A. (2015). Team adaptation: A fifteen-year synthesis (1998–2013) and framework for how this literature needs to “adapt” going forward. *European Journal of Work and Organizational Psychology, 24*(5), 652–677. https://doi.org/10.1080/1359432X.2014.1001376

McDonald, D.M., & Weir, G.R. (2006). Developing a conceptual model for exploring emergence. *In International Conference on Complex Systems*. https://strathprints.strath.ac.uk/id/eprint/2786

McLeod, A., Simkin, M. G., & Week, J.M. (2016). Toward the user-commitment continuum: establishing the importance of realisation. *International Journal of Learning Technology, 11*(2), 134–155. https://doi.org/10.1504/IJLT.2016.077521

Meyer, W.-U., Reisenzein, R., & Schützwohl, A. (1997). Toward a process analysis of emotions: The case of surprise. *Motivation and Emotion, 21*(3), 251–274. https://doi.org/10.1023/A:1024422330338

Miles, M. P., & Arnold, D. R. (1991). The relationship between marketing orientation and entrepreneurial orientation. *Entrepreneurship Theory and Practice*, *15*(4), 49–66. https://doi.org/10.1177/104225879101500407

Miller, M. J., & Feigh, K. M. (2019). Addressing the envisioned world problem: A case study in human spaceflight operations. *Design Science*, *5*. https://doi.org/10.1017/dsj.2019.2

Morgan, J. F., & Hancock, P. A. (2011). The effect of prior task loading on mental workload: An example of hysteresis in driving. *Human factors*, *53*(1), 75–86. https://doi.org/10.1177/0018720810393505

Mueller, S. T., Hoffman, R. R., Clancey, W., Emrey, A., & Klein, G. (2019). Explanation in human-AI systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI. *ArXiv.* https://arxiv.org/abs/1902.01876

National Transportation Safety Board. (1989). *Aircraft Accident Report Aloha Airlines Flight 243 Boeing 737-200, N73711, near Maui, Hawaii, April 28, 1988* (NTSB/AAR-89/03). https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR8903.pdf

National Transportation Safety Board. (1990). *Aircraft Accident Report United Airline Flight 232 McDonnell Douglas DC-10-10, Sioux Gateway Airport, Sioux City, Iowa, July 19,1989 (*NTSB/AAR-90/06). https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR-90-06.pdf

National Transportation Safety Board. (2010). *Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River, US Airways Flight 1549 Airbus A320-214, N106US, Weehawken, New Jersey, January 15, 2009* (NTSB/AAR-10/03). https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR1003.pdf

National Transportation Safety Board. (2016). *Aviation Accident Final Report Southwest Airlines Flight 3472 Boeing 737-7H4, N766SW, Pensacola, Florida, August 27, 2016* (DCA16FA217). https://www.skybrary.aero/bookshelf/books/5822.pdf

Nyuur, R. B., Brecic, R. and Debrah, Y. A. (2018). SME international innovation and strategic adaptiveness: The role of domestic network density, centrality and informality. *International Marketing Review, 35(*2), 280–300. http://doi.org/10.1108/IMR-11-2015-0239

Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. Y., Metalsky, G.I, & Seligman, M. E. P. (1982). The attributional style questionnaire. *Cognitive Therapy and Research, 6*, 287-299. https://doi.org/10.1007/BF01173577

Pervozvanski, A. A. (1992). Sensitivity, robustness and efficiency of adaptation. *International Journal of Adaptive Control & Signal Processing, 6*(3), 183–191. https://doi.org/10.1002/acs.4480060306

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643–675. https://doi.org/10.1037/0033-295X.106.4.643

Poppel, E. (1988). *Mindworks: Time and conscious experience.* Harcourt.

Powley, E. H. (2009) Reclaiming resilience and safety resilience activation in the critical period of crisis. *Human Relations, 62,* 1289–1326. https://doi.org/10.1177/0018726709334881

Prati, G., Pietrantoni, L., & Cicognani, E. (2010). Self-efficacy moderates the relationship between stress appraisal and quality of life among rescue workers. *Anxiety Stress Coping, 23*(4), 463–470. https://doi.org/10.1080/10615800903431699

Pruchnicki, S, Key, K., & Rao, A. (2019). *Problem solving/decision making and procedures for unexpected events: A literature review* (DOT/FAA/AM-20/01). Federal Aviation Administration, Office of Aerospace Medicine. https://rosap.ntl.bts.gov/view/dot/57074

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612–624. https://doi.org/10.1037/0021-9010.85.4.612

Rankin, A., Woltjer, R., & Field, J. (2016). Sensemaking following surprise in the cockpit—A re-framing problem. *Cognition, Technology & Work, 18*(4), 623–642. https://doi.org/10.1007/s10111-016-0390-2

Reason J. (1990). *Human Error*. Cambridge: Cambridge University Press.

Redish, A. D., & Mizumori, S. J. Y. (2015). Memory and decision making. *Neurobiology of Learning and Memory*, *117*, 1–3. https://doi.org/10.1016/j.nlm.2014.08.014

Reitan, R.M. and Wolfson, D. (1993) *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation (*2nd Edition). Neuropsychology Press.

Reivich, K. J., Seligman, M. E. P., & McBride, S. (2011). Master resilience training in the U.S. Army. *American Psychologist, 66*(1), 25–34. https://doi.org/10.1037/a0021897

Rivera, J., Talone, A. B., Boesser, C. T., Jentsch, F., & Yeh, M. (2014). Startle and surprise on the flight deck: Similarities, differences, and prevalence. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 58*(1). 1047–1051. SAGE Publications. https://doi.org/10.1177/1541931214581219

Robertson, I. T., Cooper, C. L., Sarkar, M., & Curran, T. (2015). Resilience training in the workplace from 2003 to 2014: A systematic review. *Journal of Occupational and Organizational Psychology, 88*(3), 533–562. https://doi.org/10.1111/joop.12120

Ross, B. H., & Spalding, T. L. (1991). *Some influences of instance comparisons on concept formation*. (pp. 207–236). Morgan Kaufmann Publishers Inc.

Sánchez-Escribano, M. G., & Sanz, R. (2014). Emotions and the Engineering of Adaptiveness in Complex Systems. *Procedia Computer Science*, *28*, 473–480. https://doi.org/10.1016/j.procs.2014.03.058

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, *37*(1), 5–19. https://doi.org/10.1518/001872095779049516

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*(4), 225–260. https://doi.org/10.1037/h0076770

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schwarz, M., Kallus, K. W., & Gaisbachgrabner, K. (2016). Safety culture, resilient behavior, and stress in air traffic management. *Aviation Psychology and Applied Human Factors, 6*(1), 12–23. https://doi.org/10.1027/2192-0923/a000091

Sheffi, Y. (2005). *The resilient enterprise: Overcoming vulnerability for competitive advantage*. MIT Press.

Shiau, W. L., & George, J. F. (2014). A grounded theory approach to information technology adoption. *Communications of the Association for Information Systems, 34*(1), 81. https://doi.org/10.17705/1CAIS.03481

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, *37*(1), 137–148. https://doi.org/10.1518/001872095779049444

Smith-Jentsch, K. A., Salas, E., & Baker, D. P. (1996). Training team performance-related assertiveness. *Personnel Psychology, 49*(4), 909–936. https://doi.org/10.1111/j.1744-6570.1996.tb02454.x

Spinka, M., Newberry, R. C., & Bekoff, M. (2001). Mammalian play: Training for the unexpected. *The Quarterly Review of Biology*, *76*(2), 141–168. https://doi.org/10.1086/393866

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1995). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In *Constructivism in Education*. Routledge.

Spiro, R. J. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains.* Center for the Study of Reading Technical Report (No. 441).

Sugimoto, I., & Murai, K. (2018). Study of instructor support system for simulator-based exercise using mental workload monitoring system. *2018 World Automation Congress (WAC)*, 1–5. https://doi.org/10.23919/WAC.2018.8430488

Sutcliffe, K. M., & Vogus, T. J. (2003) Organizing for resilience. In K.S. Cameron, J. E. Dutton, & R. E Quinn (Eds.), *Positive organizational scholarship: Foundations of a new discipline* (pp. 94–110). Berrett-Koehler.

Szalma, J. L., & Hancock, P. A. (2011). Noise effects on human performance: A meta-analytic synthesis. *Psychological Bulletin, 137*(4), 682–707. https://doi.org/10.1037/a0023987

Talone, A. B., Rivera, J., Jimenez, C., & Jentsch, F. (2015). Evaluating startle, surprise, and distraction: An analysis of aircraft incident and accident reports. *18th International Symposium on Aviation Psychology,* 278–283. https://corescholar.libraries.wright.edu/isap_2015/60

United Kingdom Civil Aviation Authority. (2013). *1036-Global Fatal Accident Review 2002 to 2011*. UK Civil Aviation Authority. https://www.caa.co.uk/Data-and-analysis/Safety-and-security/Analysis-reports/Global-fatal-accident-review/

Mjelde, F. V., Smith, K., Lunde, P., & Espevik, R. (2016). Military teams—A demand for resilience. *Work (Reading, Mass.)*, *54*(2), 283–294. https://doi.org/10.3233/WOR-162298

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, *35*(1), 205–211. https://doi.org/10.1016/s0896-6273(02)00746-8

Wells, A. (2002). GAD, metacognition, and mindfulness: An information processing analysis. *Clinical Psychology: Science and Practice, 9*(1), 95–100. https://doi.org/10.1093/clipsy.9.1.95

West, B. J., Patera, J. L. and Carsten, M. K. (2009) Team level positivity: Investigating positive psychological capacities and team level outcomes. *Journal of Organizational Behavior, 30*, 249–267. http://dx.doi.org/10.1002/job.593

Wickens, C. D. (2001). Attention to safety and the psychology of surprise. Keynote Address at the 11th International Symposium on Aviation Psychology (pp. 1–11). The Ohio State University.

Wiener, E. L. (1985). Beyond the sterile cockpit. *Human factors*, *27*(1), 75–90. https://doi.org/10.1177/001872088502700107

Wilson, R. C., Christensen, P. R., Merrifield, P. R., & Guilford, J. P. (1975). *Alternate Uses Test.* Sheridan Psychological Company.

Wing, D. J., Chancey, E. T., Politowicz, M. S., & Ballin, M. G. (2020). Achieving resilient in-flight performance for advanced air mobility through simplified vehicle operations (Document ID 20205000771). *AIAA AVIATION 2020 FORUM* (p. 2915). https://ntrs.nasa.gov/citations/20205000771

Woods, D. D. (2006a). Essential characteristics of resilience for organizations. In E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience engineering: Concepts and precepts*. Ashgate Publishing Ltd.

Woods, D. D. (2006b). How to design a safety organization: Test case for resilience engineering. In E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), Resilience engineering: concepts and precepts (pp. 315–324). Ashgate Publishing Ltd.

Woods, D. D. (2015). Four concepts for resilience and their implications for systems safety in the face of complexity. *Reliability Engineering and System Safety, 141*, 5–9. https://doi.org/10.1016/j.ress.2015.03.018

Yeomans, J. S., Li, L., Scott, B. W., & Frankland, P. W. (2002). Tactile, acoustic and vestibular systems sum to elicit the startle reflex. *Neuroscience and Biobehavioral Reviews, 26*(1), 1–11. https://doi.org/10.1016/S0149-7634(01)00057-4

Xu, Y., & Chen, Z. (2006). Relevance judgment: What do information users consider beyond topicality? *Journal of the American Society for Information Science and Technology, 57*(7), 961–973. https://doi.org/10.1002/asi.20361

Zuiderwijk, D., van der Vorm, J., van der Beek, F. A., & Veldhuis, G. J. (2016). Reinforcing resilience under changing conditions: operational resilience as an organisational power: an introduction.

Appendix A: Tables of Definitions

Definitions and features of constructs proposed to be considered in this research.

We are not attempting to find the "correct" or "perfect" definition of the key terms in this work. The purpose of this section is to ensure that (a) our team has a mutual understanding of the constructs we are researching, and (b) we provide a definition and/or description of the construct that is clear to the reader and end user of our product.

With these goals in mind, we compiled the definitions and features, selected from published research and reports, of the constructs pertinent to this work. The table that contains definitions and complementary information relating to the terms used to describe the construct in the literature with citations for each entry. The information in each cell is prefaced with a label or multiple labels that reflect the type of information entered. These labels are:

Definition – specified as a definition by the author.

Feature – is an aspect of the construct that is not specifically called-out as a definition.

Tool – contains information that includes guidance or training regarding the construct.

- *Measure* – refers to a scale or instrument used in studying the construct
- *Training* – an aspect of the information pertains to education or training
- *Note* – general aspect of the construct or term used.
- *Compare* – a comparison of two terms
- *Contrast* – the contrast between two terms

In addition, each label will specify the following if applicable:

- *Human* – pertains specifically to the human
- *Team* – pertains specifically to a team or teams
- *System* – pertains specifically to a system

It should be emphasized that the constructs described in the section are all interrelated. For example, resilience is a construct that is comprised of many different terms such as adaptability, cognitive flexibility, performance adaptation, metacognition, etc. Thus, some of the information will overlap. In two cases, we present a chart to compare and contrast constructs that are particularly intertwined, that of resilience and adaptability, and startle and surprise.

The literature search used to create our definitions is not exhaustive. However, we included major historical and modern research studies as well as the purpose of our study in our preparation. The full citations for the charted information are found in the Reference section of this document.

A.1 Definitions and Features of Resilience

Definition/Information	Source
(*Definition – System*) "The capacity of a system to exhibit a new state of operational stability when adaptation to recover the prior base state has now failed." (*Feature – System*) "Resilience is the capacity to dynamically achieve a 'new-normal.'"	Hancock & Cruit (2020)
(*Definition – Human, System*) "Resilience is an expression of how people and organizations cope with everyday situations by adjusting their performance to the outcomes." (*Feature – Human*) "Resilient behaviors are associated with learning, responding, monitoring, and anticipating." (*Feature – Human*) Resiliency is adaptive expertise.	Chandra et al. (2020)
(*Definition – System*) Resilience is the systemic capacity to change as a result of circumstances that push the system beyond the boundaries of its competence envelope. The system may have to amend some, or even all of its goals, procedures, resources, roles, or responsibilities. As a result of those changes, the work system then expresses a revised competence envelope. In effect, it becomes a different system.	Hoffman & Hancock (2017)
(*Feature – System*) Resilience results when a system is pushed beyond its adaptive capabilities. The system can no longer cope "via its traditional forms of response" and is "pushed to enact new and innovative paths toward a new behavioral stability." (*Feature – Human, System*) Resilience requires an individual or system "to be able to adapt to changing circumstances by attaining a differing form of operational stability."	Hoffman & Hancock (2017)
(*Definition – Human*) "In an important psychological sense, resilience represents how people respond to stress or trauma."	Infurna & Luther (2016)
(*Feature – Human*) Resilience is an element of personality.	Matthews et al. (2017)
(*Definition*) Resilience is "an ongoing and adaptable process including situation assessment, self-review, decision and action."	EASA (2017)
(*Definition*) The capability of an element to return to a stable state after a disruption.	Barnett & Pratt (2000) Callister & Rethwisch (2007) Dekker (2012) Hamel & Valikangas (2003) Hollnagel et al. (2006) Powley (2009) Sheffi (2005) Walker et al. (2002)
(*Definition – Human*) The ability to anticipate and adapt to the potential for surprise and failure.	Hollnagel et al. (2006)

(*Feature – System*) An unexpected or surprising event pushes a system closer to its boundaries of safe operation. If the system does not know how to respond outside those boundaries, the system is brittle and adverse outcomes can result. A system that knows how to operate outside of the safe boundaries is resilient. (*Definition – System*) Resilience in its truest sense is the system's ability to return itself to a stable system when, due to disruptions that are either foreseen or not, it operates outside of its normal capabilities as a whole.	Dekker & Pruchnicki (2013) Hollnagel et al. (2008) Nemeth et al. (2009)
(*Feature*) Resiliency as a way to prevent adverse outcomes.	Holbrook et al. (2019)
(*Feature – Human*) Resilience engineering is a paradigm that focuses on how people cope with complexity under pressure to achieve success	Woods (2006)
(*Feature – Team*) Similar to individual resilience, team resilience has also been conceptualized as both a capacity and a demonstration or team process.	Sutcliffe & Vogus (2003) West et al. (2009)
(*Feature – Team*) We view team resilience as a shared belief held by the team that it can respond to disruptive and challenging events, recover from setbacks, and thrive as a team under these conditions. (*Definition – Team*) Therefore, rather than a capacity or a demonstration of such a capacity, we view team resilience as an emergent state or one of the "*cognitive, motivational, and affective states of teams*"	Kennedy et al. (2016)
(*Note – Human*) Psychological, physical, and social factors contribute to individual resilience.	Alliger (2015)
(*Definition – Team*) We define team resilience as the capacity of a team to withstand and overcome stressors in a manner that enables sustained performance (*Feature – Team*) It helps teams handle and bounce back from challenges that can endanger their cohesiveness and performance.	Alliger (2015)
(*Feature*) Resiliency as a personal characteristic. The ability to adapt to, and to bounce back from, unexpected events (an idea of effectively returning to baseline/equilibrium).	Matthews et al. (2017)
(*Definition*) "Resilience is a complex personal characteristic with multiple neural and psychological roots."	Matthews et al. (2017)
(*Note – Human*) Training to help flightcrew improve on a failed situation and prepare them for future similar situations is good. However, training is needed to prepare flightcrew for *unknown unknowns*. This will help teams achieve future resilience. Aviation safety now requires pilots to be trained to handle unpredictability strategically	Martin (2019)
(*Definition – Human*) Resilience is the ability to adapt to environments where one is confronted with information ambiguity, incoherence, resistance and hardship.	Mjelde et al. (2016)
(*Definition*) The ability to persist in the face of challenges and bounce back from adversity.	Reivich et al. (2011)

(Definition – Team) Maynard et al. (2015) suggests that team adaptive outcomes can include not only personal health outcomes but also team performance and emergent states such as resilience, team cognition, cohesion, collective efficacy, and trust.	Maynard et al. (2015)
(Definition – Team) The relationship between team adaptation and resilience is reciprocal in nature, whereby a team that adapts in the face of a disruption is apt to enhance the team's feelings regarding resilience, and possessing such resilience is likely to set the team up for better adaptation in the face of future triggers.	Hollnagel (2017)
(Feature – System) Robustness, which is the ability of the system to operate within its normal operating boundaries when it is perturbed	Gluck et al. (2012) Hoffman & Hancock (2017)
(Definition – System) resiliency is the "capacity of a system to absorb disturbances and adapt to a change in order to retain the same function, structure, and identity"	Faerevaag et al. (2018)
(Feature – Team) Team adaptation has been viewed by some as an inherent capability of certain teams	Gibson & Birkinshaw (2004)
(Definition – System) resilience is "an ongoing and adaptable process including situation assessment, self-review, decision and action".	European Aviation Safety Association (2017)
(Definition – System) Adaptiveness, which is a system's ability to create novel solutions to problems within the normal operating boundaries	Bhamra et al. (2011) Dalziell & McManus (2004) Gallopin (2006)
(Definition – Team) Team adaptation as a process of adjusting strategies and behaviors within a team	Cannon-Bowers et al. (1995)

A.2 Definitions and Features of Adaptiveness

Definition/Information	**Source**
(*Feature*) Experts know what *to do* as well as what *not* to do in a situation. In addition, *adaptive experts* not only use what they know, they monitor their current level of understanding of a situation, continue to learn, and strive to move to a higher level of functioning. They use each new situation as a challenge which provides a forum to facilitate additional expertise.	Kochan (2005)
(*Definition*) "Adaptive experts are able to apply knowledge effectively to novel or atypical situations."	Hatano & Inagaki (1986)
(*Feature – Team*) "Team adaptive outcomes can include not only personal health outcomes but also team performance and emergent states such as resilience, team cognition, cohesion, collective efficacy, and trust."	Kennedy et al. (2016)

(*Feature – Human*) Pilots should develop, "adaptive expertise" to handle complexity which could lead to stress, which could lead to error. (*Feature – Human*) "…pilots with adaptive expertise will have more capacity to analyze and respond to any operational situation in real-time." (*Feature – Human*) "Adaptive experts are flexible" and "innovative." (*Feature – Human*) Adaptive experts will be more capable of applying their knowledge to unexpected events and may be able to better manage stress and reduce errors as a result."	Chandra et al. (2020)
(*Definition – System*) "Adaptivity is the capacity of a work system to achieve its goals, despite the emergence of circumstances that perturb it from a predetermined course by pushing it toward the boundaries of its competence envelope. (*Feature* – System) The work system is able to employ multiple resources (personnel, materiel, finance, etc.) in multiple ways to recover, or it may even develop new ways to succeed, and in doing so can move seamlessly among them. The work system can reallocate and redirect its resources and activities to retrench from the boundaries of its competence envelope and thus still achieve its goals." (*Feature – System*) "A work system is able to re-achieve its goals and continue to exercise its ordinary procedures. This theme is the restabilization notion"	Hoffman & Hancock (2017)
(*Definition*) "Adaptability is being able to strategically return to an original state after an unexpected event."	Hancock & Cruit (2020)
(*Definition – System*) "Adaptiveness, which is a system's ability to create novel solutions to problems within the normal operating boundaries."	Bhamra et al. (2011) Dalziell & McManus (2004) Gallopin (2006)
(*Feature – Team*) Team adaptation has been viewed by some as an inherent capability of certain teams	Gibson & Birkinshaw (2004)
(*Definition*) "Strategic adaptiveness refers to the ability of firms to strategically respond to challenges or crises caused by environmental turbulence."	Miles and Arnold (1991)
(*Feature*) Hardiness paired with adaptiveness; hardiness: the ability to endure difficult conditions.	Matthews et al. (2017)
(*Definition – System*) "System adaptivity is the capacity of performing structural change in order to fulfill fixed or changing requirements and tuning to new operating circumstances during the lifetime of the system."	Sánchez-Escribano & Sanz (2014)

(*Definition – Human*) Defined adaptivity as involving the following capacities: creative problem solving, coping with uncertainty, learning new tasks and skills, adapting to teamwork and collaboration, changing procedures and developing new procedures, and adapting across cultures.	Ilgen & Pulakos, (1999) Pulakos et al. (2000)
(*Definition – Team*) Team adaptation is a process of adjusting strategies and behaviors within a team	Cannon-Bowers et a. (1995)

A.3 Compare and Contrast Resilience and Adaptability

Compare or Contrast	**Source**
(*Contrast*) “In a general sense, we can view adaptability as the search for stability in an already occupied parameter space. Resilience in contrast is the achievement of a new state of stability in a different parameter space.”	Hoffman & Hancock (2017)
(*Contrast*) The main difference is the end goal; the goal of adaptation is to return to its original state, the goal of resilience is to dynamically achieve a ‘new-normal.’”	Hancock & Cruit (2020)
(*Compare – Team*) The relationship between team adaptation and resilience is reciprocal in nature, whereby a team that adapts in the face of a disruption is apt to enhance the team’s feelings regarding resilience, and possessing such resilience is likely to set the team up for better adaptation in the face of future triggers.	Hollnagel (2017)
(*Contrast*) “In a general sense, we can view adaptability as the search for stability in an already occupied parameter space. Resilience in contrast is the achievement of a new state of stability in a different parameter space.”	Hoffman & Hancock (2017)
(*Contrast*) The relationship between team adaptation and resilience is reciprocal in nature, whereby a team that adapts in the face of a disruption is apt to enhance the team’s feelings regarding resilience, and possessing such resilience is likely to set the team up for better adaptation in the face of future triggers.	Kennedy et al. (2016)

A.4 Definitions and Features of Cognitive Flexibility

Definition/Information	**Source**
(*Definition – Human*) “Cognitive Flexibility is the essential ability to assess and adapt ongoing psychological operations and to coordinate the allocation of cognitive processes appropriately in dynamic decision-making environments” This paper suggests distributed neural networks related to cognitive flexibility can enhanced by “real-time strategy game training.”	Glass et al. (2013)

(*Definition – Human*) Cognitive flexibility "refers to a person's (a) awareness that in any given situation there are options and alternatives available (b) willingness to be flexible and adapt to the situation, and (c) self-efficacy in being flexible" (*Feature – Human*) The ability to be cognitively flexible dependents on the ability to see multiple alternatives to a situation (*Feature – Human*) Willingness to be flexible may be an internal motivational state (*Feature – Human*) An individual also needs to be confident that the alternative choice will result in the desired outcome (*Tool*) This paper reports a cognitive flexibility scale	Martin & Rubin (1995)
(*Tool*) Performed validity scales for Martin & Rubin cognitive flexibility scale.	Martin & Anderson (1998)
(*Definition*) The ability to anticipate and adapt to the potential for surprise and failure Three facets of cognitive flexibility: • Ability to see complex situations as controllable • Ability to see multiple alternative explanations for events • Ability to create alternatives	Dennis & Vander Wal (2010)
(*Measure*) Performance-based measures of cognitive flexibility • Stroop color and word test • Trail Making Test (TMT) • Wisconsin Card Sorting Test (WCST)	Berg (1948) Golden (1975) Reitan & Wolfson (1993)
(*Measure*) Self-report measures of cognitive flexibility • Alternate Uses Test • Attributional Style Questionnaire (AQS) – focus on attributing causes to good and bad events • Cognitive Flexibility Scale (CFS) – focus on communication	Martin & Rubin (1995) Peterson et al. (1982) Wilson et al. (1975)
(*Definition – Human*) "Mental flexibility is 'the human ability to adapt cognitive processing strategies to new and unexpected conditions in the environment'	Canas et al. (2005) Faerevaag et al. (2018)
(*Feature – Training*) There is no set procedure for mental flexibility because such training focuses on preparing pilots to respond to unanticipated situations that have not been trained for.	Faerevaag et al. (2018)
(*Note – Human*) Mental flexibility can hinder performance when a pilot needs to return from new procedures to standard procedures	Faerevaag et al. (2018)
(*Definition*) Cognitive flexibility is the "ability to spontaneously restructure one's knowledge in many ways, in adaptive response to radically changing situational demands.	Spiro et al. (1995)

A.5 Definitions and Features of Metacognition

Definition/Information	**Source**
(*Definition)* Metacognition is the ability to monitor one's current level of understanding and decide when it is and when it is not adequate.	Kochan (2005)
(*Definition*) Metacognition is defined as the ability to reflect upon, understand, and control one's learning. (*Training*) Metacognitive activities that help control one's thinking, or learning is what is referred to as regulation of cognition, and usually includes three essential skills: planning (strategy and resource allocation), monitoring (awareness), and evaluation (appraising the products of learning).	Schraw & Dennison (1994)

A.6 Definitions and Features of Decision-Making (Judgment)

Definition/Information	**Source**
(*Definition*) Aeronautical decision-making (ADM) is decision-making in a unique environment—aviation. It is a systematic approach to the mental process used by pilots to consistently determine the best course of action in response to a given set of circumstances. It is what a pilot intends to do based on the latest information they have.	FAA-H-8083-25B FAA-H-8083-2
(*Definition*) We define decision as the mental processes (cognitive process) that results in the selection of an action."	Shiau & George (2014)
(*Definition – Human*) "We define decision-making as the process of selecting an action."	Redish & Mizumori (2015)
(*Definition – Human*) Rational Judgment is the ability to discover and establish the relevance of all available information relating to problems of flight, to diagnose these problems, to specify alternative courses of action, and to assess the risk associated with each alternative. (*Definition – Human*) Motivational Judgment is the motivation to choose and execute a suitable course of action within the available time frame. Where the choice could be either action or no action and "suitable" is a choice consistent with "societal" norms.	Jensen (1995)
(*Definition*) Naturalistic Decision Making is the way people use their experience to make decisions in field settings.	Klein (1993)
(*Note*) It appears the expert's ability hinges on the recognition of patterns and consistencies that clarify options in complex situations. (*Feature*) Experts appear to make provisional sense of a situation, without actually reaching a decision, by launching experience-based actions that in turn trigger creative revisions.	FAA-H-8083-25B

(*Definition*) "We define decision making as a psychological process in which the decision maker chooses between various alternatives with the intent of reaching a maximal number of goals, while avoiding damage and unnecessary risks, and by using a minimal amount of resources."	Cohen (2008) Heichal (1992)

A.7 Definitions and Features of Recognition

Definition/Information	**Source**
(*Definition)* "We define 'recognition' as the act of identifying something that has been previously seen, heard or known."	McLeod et al. (2016)
(*Definition*) "In this paper we define recognition as realizing something totally with sense, perception, awareness, or behavior."	Sugimoto & Murai (2019)
(*Definition*) Recognition-primed decision making is how people come to a rapid decision when faced with a complex situation in the real-world.	Klein (1993)

A.8 Definitions and Features of Novelty

Definition/Information	**Source**
(*Definition)* "We define novelty as an individual's tendency to be drawn toward new experiences."	Gordon & Luo (2011)
(*Definition*) "Psychology researchers define novelty as a property of a stimulus that has not been previously presented to or observed by and is thus unfamiliar to the subject."	Xu & Chen (2006)
(*Definition*) "We define novelty as the emergence of a sustainable new entity with distinctly different interaction patterns."	McDonald & Weir (2006)
(*Definition*) "We define novelty as the appearance of a new sustainable entity, with interaction models which are distinct and different from its predecessors."	Dumitrescu et al. (2017)

A.9 Definitions and Features of Startle

Definition/Information	**Source**
(*Definition*) The startle reflex is an autonomic response to an unexpected auditory, visual, or tactile stimulus with an abrupt onset. (*Feature*) The physical reflexive startle response begins with muscle contractions (eye-blinks, head ducks, and crouched shoulders) followed by quick movement away from the stimulus. (*Note*) The startle response includes the startle reflex and emotional and cognitive responses where attentional resources are oriented toward the startling stimulus.	Davis (1984)

(*Feature*) Startle can be measured in a person's frequency of eye blinks, in the contraction of their neck and facial muscles, the termination of ongoing behaviors, an increase in physiological arousal, and the presence of fear or anger. (*Feature*) The startle reaction is brief and physiological. (*Feature*) Startle is a reaction to a sudden, intense, or threatening stimulus (*Note*) Startle causes closer examination of stimulus which can increase stress.	Landman et al. (2017)

A.10 Definitions and Features of Surprise

Definition/Information	**Source**
(*Definition*) Surprise is an emotional and cognitive response to unexpected and difficult to explain events. These events force people to change their understanding of the situation.	Landman et al. (2017)
(*Feature*) Managing surprise requires effective sensemaking activities. Such activities involve the search for a frame of the situation. This frame is used to focus attention, interpret information, and make sense of the situation. (*Feature*) Surprise occurs frequently in aviation but often remain inconsequential (*Feature*) An example of surprise is aviation is a subtle technical failure.	Landman et al. (2017)
(*Feature*) Surprise can be insidious; surprise can be subliminal; while the unusual is generally unexpected and surprising, what is unexpected and surprising to us is typically something common, trivial, and mundane; and more often than not, there were cues available to suggest that the unexpected should have been expected. (*Feature*) Surprise can be elicited by the presence, but also by the absence of stimuli.	Kochan & Breiter (2005)
(*Definition*) The psychology of surprise is about how people respond to unexpected events.	Wickens (2001)
(*Feature*) Surprise results from a disparity between a person's expectations and what is actually perceived.	Horstmann (2006)

A.11 Compare and Contrast Startle and Surprise

Compare/Contrast	**Source**
(*Compare*) By definition, unexpected events and/or intense stimuli always cause surprise and/or startle. (*Contrast*) Not all unexpected events lead to large physiological and emotional reactions.	Fields (2020)

(*Compare*) "Like startle, (surprise) may cause acute stress." (*Contrast*) Surprise can occur in the absence of startle. (*Contrast*) Startle is a different response from surprise with different causes, although they are often and incorrectly used interchangeably.	Landman et al. (2017)
(*Contrast*) Surprise is an emotional response to an unexpected event. Example: The airplane stalls when on autopilot at high altitude owing to insufficient thrust available to maintain airspeed while keeping a level altitude, as the pilot had become distracted and was not able to monitor the flight properly whereas startle is a reflexive response to an unexpected event. Example: a pilot winces and blinks after a windshield suddenly cracks in flight.	Kochan (2005)
(*Contrast*) Surprise "increases arousal and draws attention to its cause in an orienting manner and less in a defensive or 'flinching' manner."	Bradley (2009)

A.12 Definitions and Features of Stress

Definition/Information	**Source**
(*Feature*) Look to decrease stress in order to increase safety in aviation. Increased operational complexity associated with performance-based navigation can lead to anxiety and the stress response that results can weaken human performance in different ways by disrupting cognitive structures and processes.	Chandra et al. (2020)
(*Definition*) Stress is a response to threatening situations that involves biological, cognitive, behavioral, and emotional components.	Dismukes et al. (2015)

A.13 Definitions and Features of Safety

Definition/Information	**Source**
(*Definition*) Safety I believes systems need protection from humans and more automation to maintain safety. (Definition) Safety II considers human-skill to be valuable in maintaining safety by managing normal and unexpected situations.	Schwarz et al. (2016)
(*Definition*) "The state in which risks associated with aviation activities, related to, or in direct support of the operation of aircraft, are reduced and controlled to an acceptable level."	International Civil Aviation Organization (ICAO; 2018)

A.14 Definitions and Features of Monitoring

Definition/Information	Source
(*Feature – Team*) Monitoring is an active on-going process, typically reserved for tasks where pilots are responsible for the end result.	Chandra (2017)
(*Definition – Team*) Each flightcrew member must carefully observe the aircraft's flight path and systems and actively cross-check the actions of other crewmembers. (*Feature – Team*) Effective monitoring and cross-checking can be the last line of defense that prevents an accident because detecting an error or unsafe situation may break the chain of events leading to an accident.	FAA AC 120-51E
(*Note – Team*) Pilot Monitoring (PM) monitors the aircraft state and system status, calls out any perceived or potential deviations from the intended flightpath, and intervenes if necessary	FAA AC 120-71B
(*Definition – Human*) Effective Monitoring. A pilot is effectively monitoring if they are: 1. Following SOPs consistently; 2. Clearly communicating deviations to other crewmembers; 3. Effectively managing distractions; 4. Remaining vigilant; 5. Advising the PF if the flight guidance modes or aircraft actions do not agree with expected or desired actions and intervening if necessary; 6. Continuously comparing known pitch/power settings to current flightpath performance; and 7. Considering that the primary flight displays (PFD), navigation displays (ND), and other sources of information (for example, electronic flight bag (EFB)), might be displaying incorrect information and always on the lookout for other evidence that confirms or disconfirms the information the displays are providing.	FAA AC 120-71B

A.15 Definitions and Features of Robustness

Definition/Information	Source
(*Definition – System*) Robustness is the ability of the system to operate within its normal operating boundaries when it is perturbed.	Gluck et al. (2012) Hoffman & Hancock (2017)
(*Definition – System*) "Robustness refers to the ability of a system to respond adequately under unanticipated runtime conditions."	Sánchez-Escribano & Sanz (2014)
(*Definition – System*) Robustness is the ability of a work system to maintain effectiveness across a range of tasks, situations, and contexts.	Hoffman & Hancock, (2017)
(*Definition – System*) Robustness is conceptually defined as the ability of a work system to "maintain its function when it is perturbed"	Gluck et al. (2012) Hoffman & Hancock (2017)

A.16 Definitions and Features of Brittleness

Definition/Information	Source
(*Definition*) Brittleness, descriptively, is a rapid fall off or collapse of performance that occurs when events push a system beyond its boundaries for handling changing disturbances and variations. (*Note*) Resilience is the opposite of brittleness. Since the word resilience is used in many different ways, a new term was needed to refer to system characteristics that overcome the risk of brittleness-induced failures — Graceful Extensibility.	Woods (2015)
(*Note*) Identify areas of brittleness to anticipate possible points of system failure and thus identify appropriate interventions.	Gomes et al. (2009)

A.17 Definitions and Features of Unexpected Events

Definition/Information	Source
(*Definition*) An unexpected event is (a) An event incongruent with expectations as determined by base rate probabilities (average probability of event occurring) and the contextual information available; may be normal, abnormal, or emergency in nature; it may also be frequent, infrequent, or novel, or (b) the absence of an expected event.	Kochan & Jentsch (2003)
(*Definition*) By definition, unexpected events and/or intense stimuli always cause surprise and/or startle. (*Feature*) Not all unexpected events lead to large physiological and emotional reactions.	Fields (2020)
(*Definition*) An unexpected event as any event that takes someone by surprise, which can violate a pilot's expectations and can affect the mental processes used to respond to the event.	FAA AC 120-109A (2015)
(*Feature*) When faced with an unexpected event, pilots must evaluate the operational circumstances and correctly respond, irrespective of the nature of the problem (e.g., mechanical malfunction, environmental factors, or security). (*Feature*) Unexpected events can involve competing information and signal overload (which increases stress and cognitive workload) and can be ill-defined. Flightcrews also experience constraints like time pressure, cognitive limitations, and teamwork/communication demands. Their ability to successfully respond to the unexpected event will decline as the complexity and time pressures increase.	Pruchnicki (2020)

Appendix B: List of Acronyms Used

ADM	Aeronautical Decision-Making
AOC	Airline Operations Center
AQS	Attributional Style Questionnaire
AI	Artificial Intelligence
CAA	Civil Aviation Authority
CFS	Cognitive Flexibility Scale
CRM	Crew Resource Management
EFB	Electronic Flight Bag
EASA	European Aviation Safety Agency
EASP	European Aviation Safety Plan
EPAS	European Plan for Aviation Safety
FAA	Federal Aviation Administration
FSTD	Flight Simulation Training Device
IFP(s)	Instrument Flight Procedure(s)
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
LOC-I	Loss of Control Inflight
NTSB	National Transportation Safety Board
ND	Navigation Display
ORM	Operational Risk Management
PBN	Performance Based Navigation
PM	Pilot Monitoring
PFD	Primary Flight Displays
RRM	Risk and Resource Management
TEM	Threat and Error Management
TMT	Trail Making Test
UAS	Unmanned Aircraft System
WCST	Wisconsin Card Sorting Test