ADDRESSING THE CHALLENGES IN NEXTGEN DECISION MAKING

Decision Making Capabilities and Limitations, Topic 12-04

Kathleen Mosier, Co-Investigator

San Francisco State University

Ute Fischer, Co-Investigator

Georgia Institute of Technology

Judith Orasanu, Principal Investigator

NASA Ames Research Center

Interim Report #3
FAA/NASA NextGen Flight Deck Human Factors Research
Interagency Agreement #DTFAWA-10-X-80005, Annex 9

Submitted to Dr. Tom McCloy
Scientific and Technical Advisor for Human Factors
FAA Human Factors Research and Engineering, ATO-P

May 14, 2012

POC: Judith Orasanu, NASA Ames Research Center, Mail Stop 262-4, Moffett Field, CA 94035-0001, Judith.Orasanu@nasa.gov, *650-604-3404*

Table of Contents

ADDRESSING THE CHALLENGES IN NEXTGEN DECISION MAKING

The Challenges

In *Flight Crew Decision Making: Now and NextGen* (Mosier, Fischer, & Orasanu, 2011), we presented the components of the decision-making process and some of the most important and relevant models of decision making; described what experienced crews and automation bring to the decision process; defined challenges and impediments to decision making; and discussed challenges for decision making in NextGen operations. These challenges included:

Information overload - The technological capability to provide almost unlimited information to flight crews increases the possibility of information overload. Pilots will be challenged to define and access needed information in a timely manner, and to integrate information from several sources to create a coherent situation model. Depending on information design and display characteristics, the ability to pattern-match may be reduced, and more analytical processing may be required – with its attendant increase in cognitive demands.

Development of new expertise - In the early stages of NextGen, flight crews will be unfamiliar with new systems and procedures, and will need to develop expertise in the changed operational domain. The knowledge store that formed the basis for informed heuristics will need to be updated, and the use of shortcuts that work in current operations may lead to errors or mis-assessments in NextGen operations. For example, crews may work from old models rather than consider new automation capabilities when judging the feasibility of a low-visibility CSPR approach, or may break out of a paired approach based on old standards of safe separation.

Overreliance on automation - Because so much of the flying task will be accomplished via automated systems, the tendency to rely on them in decision-making tasks may be exacerbated, especially under time pressure. Susceptibility to automation bias may increase. Flight crews will be dependent on automation for maintaining efficient TBO, and may be reluctant to question automation functioning or to make tactical interventions if it means this efficiency would be diminished.

Non-vigilance and fatigue - Extensive research has demonstrated that humans are not good at monitoring for long periods of time, and non-vigilance and complacency are certain to be challenges in the highly automated NextGen environment. Moreover, flight crews engaged in prolonged monitoring are likely to experience fatigue.

Erosion of skills - Dependence on automation will also serve to keep pilots out of the flying-task loop, fostering complacency and an erosion of flying and situation assessment skills. This may be the most critical challenge of NextGen operations – to ensure that flight crews maintain proficiency so that they are prepared to take over from automation when needed. A test of NextGen operations will be how well they can accommodate anomalies and off-nominal events. Skilled flight crew response will be key.

In this report, we discuss training and human factors/design literature and recommendations to address these challenges. Section 1 describes interventions geared toward pilots: training for coherence, development of new expertise in situation awareness (SA) and decision making, methods for mitigating decision biases, supporting vigilance, and maintaining flying skills in the highly automated NextGen environment. We include a discussion of team issues and team training to enhance crew resource management, team SA and metacognition, and collaborative procedures. In Section 2 we turn to automation and decision support systems, and address the issue of how their design and display can facilitate coherent decision processes and the development of new expertise, mitigate decision biases, and support appropriate cognition in NextGen operations.

Definitions of Key Concepts

In Mosier et al. (2011), we offered several terms to describe judgment and decision making: front-end and back-end phases; coherence and correspondence strategies and goals; and tactics ranging from intuition to analysis. These terms are integral to this report as well, and so are defined below.

Front End and Back End Decision Making Processes

The terms front end and back end delineate two phases of decision making. Front-end processes make up what is most often referred to as the judgment phase of decision making. Related terms are diagnosis, situation assessment and awareness, building situation models, and pattern recognition. Front-end processes such as acquiring, filtering, integrating, and assessing information result in a judgment, which may be a rather straightforward evaluation of the initial cue (as when flight crews judge their fuel remaining as insufficient to reach their destination airport), or may involve a complex mental representation (as when a pilot integrates status indicators from various systems to diagnose a problem). An individual's judgment may be based on intuitive processes such as pattern recognition, or analytical processes involving deliberation. Once operators make a judgment about a problem, their judgment triggers decision processes.

Decision processes form the back end of decision making and culminate in selecting and executing a course of action. Back-end processes may involve retrieving an appropriate course of action from memory, locating a prescribed response in the appropriate manual, adapting a known response to the specific demands of the current situation, mentally simulating a possible response, planning a sequence of actions, or evaluating alternatives (Orasanu, 2011; see also Parasuraman, Sheridan, & Wickens, 2000, for a taxonomy of stages of decision making).

Applications that support aviation decision making, such as training or system design, must be approached very differently depending on the target phase. For example, decision support geared toward front-end processes will facilitate diagnoses, situation assessment and situation awareness; systems geared toward back-end processes will focus on aiding option generation, selection or choice of actions.

Correspondence and Coherence Strategies and Goals

The terms correspondence and coherence have been used to denote standards for the goodness of decision-making processes, strategies, and/or goals of decision making.

Correspondence refers to the empirical accuracy of judgment and decision making, and coherence refers to the criteria of rationality and consistency in judgment and decision processes (Hammond, 1996, 1999, 2000).

Correspondence. Correspondence, or empirical accuracy, is a function of how well an individual uses multiple and often fallible (i.e., probabilistic) indicators to make judgments and decisions in the world. For example, a pilot strives for correspondence when checking cues to judge height and distance from a runway and deciding when to turn for final approach. Features of the environment and the quality of available cues impact the accuracy of both front-end and back-end processes. Cues that have high ecologically validity, that are concrete and/or can be perceived clearly will facilitate accurate judgments and correct decisions. Cues that are ambiguous, or are murkier because they are not as concrete in nature or are obscured by factors in the environment, will hinder accurate judgments and adversely impact decisions (Mosier, 2002).

Although correspondence is most obvious as a criterion for the outcome (back-end) of decision making—clearly, accurate decisions are the goal—it is also a criterion for underlying front-end processes. It has been documented that operators need to have “good” situation awareness to successfully make decisions and perform their tasks (Endsley, 1996, 2000). However, the requirements for SA in the NextGen environment will be heavily focused on coherence criteria.

Coherence. The criterion of coherence, or rationality and consistency, is most often applied to front-end processes, and is more difficult to measure than correspondence. Coherence has traditionally been evaluated against the standard of mathematical models, and is expected to follow Bayesian (1958) laws of probability. This concept has been central in the research on heuristics and biases: the use of heuristics is viewed as a non-coherent strategy that may lead to biased judgments and decisions.

In NextGen operations, coherence must be defined differently, as deterministic information rather than probabilistic cues provides the basis for front-end processes. Coherence depends upon decision makers using all relevant information logically and consistently. A pilot, for instance, seeks coherence when configuring the aircraft for landing, ensuring that all system parameters displayed in the cockpit are appropriately set. What is important for coherence in automated environments is completeness and consistency of the judgment, diagnostic, or situation assessment process, especially in cases where empirical accuracy is not known or not easily accessible (Hammond, 1996; 2000; 2007; Mosier, 2002; 2009). To a significant degree, the use of intuition or analysis depends on the characteristics of the cues in the environment: engineered numerical or textual information typically induces analysis, whereas graphic or non-numeric information typically induces intuitive processes.

Intuition and Analysis: Systems 1 and 2

The notion of a cognitive continuum from intuition → analysis was developed by Hammond (e.g., 1993; 1996; 2000). *Analysis* refers to a "step-by-step, conscious, logically defensible process," whereas *intuition* typically describes "the opposite - a cognitive process that somehow produces an answer, solution, or idea without the use of a conscious, logically defensible, step-by-step process" (Hammond, 1996, p. 60). During the judgment process, individuals may move along this continuum, oscillating between intuition and analysis - or stopping at points on the continuum (Hamm, 1988). Pilots, for example, may use intuition when gauging weather from clouds ahead, switch to analysis to read and interpret printed weather data, and utilize some combination of the two to judge the safest path, or decide whether to continue on or turn back.

Kahneman (e.g., 2003; Kahneman, Lovallo, & Sibony, 2011) used the terms *System 1* and *System 2* to refer to intuitive and reflective (or analytical) modes of cognition. Both Hammond and Kahneman characterize the two modes as differing on several cognitive properties. System 1, or intuition is characterized by rapid data processing, with low cognitive control and little to no conscious awareness of processes. In contrast, analysis demands high cognitive control and conscious awareness. It requires task-specific organizing principles, and slow processing of data (see Hammond, McClelland, & Mumpower, 1980; Hammond, Hamm, Grassia, & Pearson, 1997).

These cognitive systems are not in themselves good or bad in terms of judgment and decision making. However, each system is associated with particular possibilities for error. System 1 cognition often involves the use of heuristics, which can lead to biases in judgments and decisions. System 2 analysis requires more time than System 1 intuition, and is brittle in that small slips in the process can produce large errors. The appropriateness of each system depends on the task context as well as the expertise of the decision maker. Naturalistic contexts or environments with perceptual, probabilistic cues may be amenable to System 1 pattern matching processes; electronic environments with data displays and layered information will require more mindful analytical processes. Additionally, expertise may enable System 1 short cuts in the naturalistic environment - expert pilots, for example, may recognize a situation as similar to past experiences, and may be able to use intuitive recognitional processes under conditions that would demand analysis of a novice. Pattern matching short cuts, however, may be ineffective or even counterproductive in the electronic environment despite high pilot expertise, as the interpretation and integration of electronic data require reflective thought.

Section 1. Crew Training: Current Operations and NextGen Recommendations

The single defining theme from the many discussions conducted around the country was that a focus on quality, not just quantity, of training and experience is essential to meaningful improvement. While participants generally agreed that experience (i.e., total flight time) can be an indicator of a pilot's proficiency and suitability for part 121/135 operations, they stressed that quality of training and quality of experience are far more important in determining an individual's readiness to operate in the air carrier environment.

Both in Washington and in cities around the country, participants in the Call to Action noted the various elements of a generational "paradigm shift" in the pilot population, which involves a fundamental shift in experience, expectations, and work practices. Participants stressed that it is critical to ensure that training is modified to address and accommodate these changes. There is, however, no consensus on what those changes should include. Some participants urged a back-to-basics focus on stick-and-rudder skills. Others argued that today's focus on automation management, CRM, decision making, and TEM is far more relevant to today's part 121 and part 135 operational environment. (FAA, 2010, pp. 22-23)

Training Situation Awareness: Coherence and Correspondence

How do we train pilots to be good decision makers in the highly automated, information rich NextGen environment? Broadly, training for decision making will need to focus on front-end processes such as situation assessment and situation awareness, and should emphasize process vigilance. Training must provide pilots with mental models of system functioning that enable them to identify all necessary information, as well as information gathering strategies that ensure that they identify, attend to and integrate all appropriate and relevant cues and information to make decisions. Although current training has given much attention to expert pattern recognition or matching, it may be lagging with respect to improving pilots' coherence skills through training and practice. Because automation will be one if not *the* defining characteristic of NextGen operations, pilots will be making most decisions using data from the highly sophisticated electronic systems, rather than probabilistic cues. Training for expertise in this environment will entail practice in coherent processing, particularly when systems and procedures are slightly different from current practice. In NextGen, for example, the increased reliance on automated systems for precise 4D trajectories, provision of more information in the cockpit, and new capabilities such as all-weather operations on Closely Spaced Parallel Runways (CSPRs) will increase the importance of coherence in decision making. Closely spaced parallel approaches will be conducted in IMC using new conformance monitoring displays and automated assistance for speed and spacing control. Pilots will be given more responsibility for spacing and will have new automated tools to help them.

Situation awareness (SA) concerns correspondence and coherence aspects of situation comprehension (e.g., Endsley, 1996). Pilots need to correctly identify and assess elements outside of the cockpit, such as other aircraft or mountains, as well as within the cockpit, such as autopilot mode or system states, and they need to integrate this information into a coherent mental representation, a situation model (Durso, Rawson, & Girotto, 2007; Orasanu, 1994).

SA involves perceptual, interpretive and knowledge integrating processes (Durso & Gronlund, 1999; Durso, et al. 2007; Endsley, 1995; 2000); however, current pilot training aimed at improving SA has focused mainly on the first two processes. This emphasis is consistent with both the notion of expert performance as pattern recognition (Klein, 1989; 1993) and analyses of pilots' SA-related errors (Wickens, 2001). Error analyses revealed that pilots frequently overlooked or misinterpreted critical cues (Jones & Endsley, 1996) or failed to notice task-relevant changes in flight conditions (Muthard & Wickens, 2002; Sarter & Woods, 1994). Problems with the integration of available information, in contrast, were less frequently observed (Jones & Endsley, 1996). Consequently, training must be geared toward improving pilots' cue recognition and awareness of threats, and providing them with strategies that support effective monitoring behavior and attention allocation (e.g., Bolstad, Endsley, Costello, & Howell, 2010; Endsley & Robertson, 2000; Salas, Cannon-Bowers, Fiore & Stout, 2001; Shebilske, Goethl & Garland, 2000).

SA training should also be cognizant of the phenomenon of change blindness, in which individuals fail to detect changes in an image or scene. This effect has been found to be extremely robust (see Levin & Simons, 1997; Rensink, 2002 for reviews). Early studies in change blindness found that participants instructed to view photos for a memory test later failed to notice when rather large changes occurred, for instance when the men in the photo changed hats (Grimes, 1996). One very common misconception about change blindness is the belief that it only occurs because the participants' attention is elsewhere (Simons, 2000). To the contrary, multiple studies have demonstrated that attending to where the change occurs is not enough for better than average detection of the change (Ballard, Hayhoe, & Pelz, 1995; Simons, 1996). What is necessary is that one attends to and encodes individual features of a scene in order to detect subsequent changes in that scene (Simons, 2000), although changes affecting the 'gist' or overall theme of an image are noticed more often than changes that keep the gist intact (Sampanes, Bridgeman, & Tseng, 2008).

The implications of change blindness for SA are evident - blindness to significant changes in system status or the operational environment can be deadly. One problem for designing training to counteract change blindness is that people are very poor at predicting their own ability to detect changes (Levin, Momen, Drivdahl, & Simons, 2000). Improved displays as well as training interventions are required to lessen its potential impact on SA. For example, cluttered displays can lead to increased change blindness (Treisman & Gelade, 1980); in contrast, displays that minimize clutter facilitate operator response, decrease the mental demands of the task, and increase operators' awareness of significant threats (St. John, Smallman, Manes, Feher, & Morrison, 2005).

Loss of SA may occur when crews are interrupted during the performance of sequenced tasks. Pilots may not remember where in the task sequence they had been prior to the disruption and which tasks they already had completed. A technical solution to this problem is provided by the CHEX system (Change History EXplicit; St. John, Smallman, & Manes, 2005), which logs relevant changes in a situation and presents them in a sortable table. CHEX has been shown to be quite effective in supporting operators' change detection and identification, especially in comparison to a system that only allows for the simple replay of events at high speeds (St. John, Smallman, & Manes, 2005).

SA is also addressed in the broader context of crew resource management training, especially as part of courses on threat and error management [see for example the Enhanced Safety through Situation Awareness Integration (ESSAI) training developed by Hörmann et al., 2003a, b; 2004]. Consistent with their overall objective to teach pilots effective error trapping strategies, courses emphasize the identification of and planning for potential threats. Pilots are instructed to look for "red flags" –cues that indicate a potential threat—to imagine worst case scenarios and to get prepared for these events. While 'thinking ahead' and 'taking preparatory steps' are cognitive activities that rest on coherence processes, training does not explicitly address these information and knowledge integrating processes. How pilots construct coherent mental models of flight situations is an issue that is rarely touched upon in training. Nor do SA or threat and error management training sufficiently acknowledge the importance of meta-cognitive processes that are necessary for pilots to monitor and assess the adequacy of their situation comprehension. On the other hand, coherence and metacognitive processes will be critical in pilot decision making in NextGen operations. As the NextGen cockpit will involve more automated systems and functions, pilots will be required to monitor and make sense of a larger, more diverse and more distributed set of system indications, heightening the need for training initiatives that target coherence processes.

Constructing coherent situation representations and metacognition are central issues in theories of naturalistic decision making (NDM). Klein, Phillips, Rall, and Peluso (2007) discuss the role of frames in decision makers' efforts to understand situations. Frames are conceived of as explanatory structures that are mental representations of individuals' experiences. Cues and information activate frames that in turn help decision makers to interpret situations by defining, connecting and filtering data. While the Klein et al.'s (2007) data-frame theory concerns deliberate sensemaking, the recognition and metacogniton (R/M) model of decision making by Cohen and colleagues (Cohen, Freeman, & Wolf, 1996; Cohen, Freeman, & Thompson, 1997) includes intuitive as well as analytic processes. The R/M model assumes that story structures guide decision makers' cue recognition and form the basis for critiquing and correcting their situation comprehension. Although both models were originally developed to account for decision making by intelligence and military personnel, they generalize to other domains, including pilot decision making (Fischer, Orasanu, & Davison, 2003; Orasanu, 2010). Further, Cohen, Freeman and Thompson (1997) designed a training program that is based on the R/M model and coaches trainees to explain their situation assessment, evaluate the reliability of their story and to consider alternative stories. This type of meta-recognitional skills training was found to improve army commanders' battlefield situation assessment. It also could provide a useful approach to teaching pilots the analytical skills necessary to cope with the informational complexity present in the NextGen flight deck.

Coherence in the electronic cockpit requires System 2 analytical processing because the data comprising the electronic 'story' are impossible to evaluate intuitively and heuristics or recognitional shortcuts may not work. Training for NextGen will need to include instruction and practice in using analytical skills to find relevant data and information, assess automated system functions and displays, and develop highly accurate mental models of how the automation works. The brittleness of automated systems must be dealt with: failures may be abrupt and unpredictable rather than preceded by graceful degradation (Billings, 1996; Woods & Sarter, 2000), and one tiny detail that is the least bit inconsistent (e.g., a decimal point in the wrong place) can destroy the coherence of the electronic story. These small discrepancies may be the most difficult to detect.

Given the character of automated aids that experts are and will be dealing with in NextGen (e.g., FAA, 2009), the importance of comprehensive training for system users cannot be overemphasized. SA training for NextGen operations will also need to strengthen pilots' system understanding to ensure adequate system monitoring and mode awareness (Lyall, Boehm-Davis, & Jentsch, 2008; Mumaw, Sarter & Wickens, 2001). SA is both data-driven (e.g., more salient cues are more readily perceived) and knowledge-dependent. However, as the environment in which operators act becomes more automated, knowledge-driven processes, and thus the quality of an individual's underlying knowledge structures, gain in importance. Much research on current operations has focused on misunderstandings of complex systems in the cockpit (such as the flight management system) that result from mismatches between the mental model of the pilot and the behaviors of the system (Sarter & Woods, 1995; Sherry & Polson, 1999; Sheridan & Parasuraman, 2006). For instance, Sarter and colleagues (Sarter, Mumaw, & Wickens, 2007) observed that pilots' failures to detect inappropriate mode annunciations resulted from inaccurate system knowledge rather than a failure to see the system indication – in fact, pilots frequently fixated the FMA when the problem occurred. Similarly, mode errors, or more generally mode confusions have been associated with gaps and/or misconceptions in pilots' mental models of flight deck automation (Mumaw, Sarter & Wickens, 2001; Plat & Amalberti, 2000; Woods & Sarter, 2000).

Billings (1996) emphasized the need to train pilots *how systems operate* rather than simply *how to operate systems*: "If a pilot does not have an adequate internal model about how the computer works when it is functioning properly, it will be far more difficult for him or her to detect a subtle failure. We cannot always predict failure modes in these more complex digital systems, so we must provide pilots with adequate understanding of how and why aircraft automation functions as it does" (p. 96). Operators must have sufficient knowledge of what new automated systems can do, what they "know," and how they function within the context of other systems, as well as knowledge of their limitations, in order to utilize them efficiently and exploit their real capabilities. These are coherence issues.

Current approaches to automation training do not provide pilots with an adequate conceptual understanding of flight deck automation (Mumaw, Boorman, & Griffin, 2001; Pariès & Amalberti, 2000; Plat & Amalberti, 2000; Woods & Sarter, 2000). Instead, training focuses on basic automation features and functions and offers trainees recipes for their use (Mumaw et al., 2001; Woods & Sarter, 2000). Fundamental properties of automation as well as strategic questions of when to use different features and functions are rarely, if at all, addressed (Mumaw et al., 2001). Recently, several training guidelines and programs have been proposed to redress these omissions. Casner (2003) conducted an automation training study involving students from a professional pilot academy. Study participants received eighteen hours of classroom instruction that emphasized fundamental cockpit automation concepts and skills. At the end of training, participants demonstrated a solid understanding of automation functions and principles, as assessed in a written exam and during a check flight in a full-mission simulator (average scores were between 82 and 91 percent).

Mumaw, Boorman and Griffin (2001) designed a computer-based automation training program to help pilots transitioning to highly automated flight decks acquire simple but effective mental models of the automation. Instruction proceeds in individual modules that form a linked hypertext network. Task modules are organized around a single flight task, provide task-relevant automation choices and explain their advantages and disadvantages. As trainees go through task modules they can follow links to mode modules and system concept modules to explore the functions, usages and underlying design principles of modes mentioned in the context of a task. The best teaching practices discussed by Lyall and colleagues (Lyall, Boehm-Davis & Jentsch 2008; Lyall, Jentsch, & Boehm-Davis, 2006) echo this approach. In addition, they stress the importance of scenario-based learning and consequently advocate that automation training is crew-based. Crew-based training has the added advantage that it supports peer learning, especially if crewmembers are paired in terms of their FMS knowledge. The most effective crew combination was found to involve a highly knowledgeable captain and a medium knowledgeable first officer. The benefits of many aspects of crew training will be further discussed in a forthcoming chapter that addresses biases and limitations of crew-level decision making.

New Tasks and New Tools: Development of Expertise for New Systems

Even the most seasoned professional pilots will experience challenges to their expertise as new automated systems and procedures are introduced in NextGen operations. The training issue for NextGen is less a matter of how to teach but rather *what* to teach to enable the development of new expertise. That is, we need to define how expert performance in NextGen will differ from expert performance in current operations. What generalizes? What current skills, knowledge, or mental models are likely to lead pilots astray because they do not apply? Where will pilots likely have skill or knowledge gaps?

Much of the applied work that has been done on expert processes in aviation has focused on intuitive and sensory-driven correspondence strategies. Klein's model of expert Recognition-Primed Decision Making (e.g., Klein, 1993, 2000; Zsambok & Klein, 1997), for example, describes expertise as the ability to identify critical cues in the environment, to recognize patterns of cues, and to understand the structural relationships among cues. According to this model, expert pilots look for familiar patterns of relevant cues, signaling situations that they have dealt with in the past, and base their responses on what they know "works" (e.g., Klein, 1993; Klein, Calderwood, & Clinton-Cirocco, 1986).

Experienced pilots typically are highly competent in these types of correspondence strategies. They are adept at assessing cue validity within specific situational contexts, and they are better than inexperienced pilots at predicting the outcome of a given decision or action as their expectations have been shaped by a wide array of experiences. Expertise offers a great advantage in correspondence judgments, as expert pilots are able to quickly recognize a situation from patterns of probabilistic cues, and may be able to use intuitive judgment processes under conditions that would demand analysis from a novice. For example, novice pilots may need to use a combination of computations and cues outside of the aircraft to figure out when to start a descent for landing. In contrast, experienced pilots may look outside the cockpit window and intuitively recognize when the situation 'looks right' to start down.

Results of a study examining performance in paired simultaneous CSPR approaches illustrate the trade-offs involved in mixing old and new procedures and automation (Verma et al., 2011). Pilots joined a paired approach at a coupling point and continued to SFO; they were responsible for maintaining a 15sec temporal separation from the lead aircraft. Display type – position vs. prediction display – and current automation vs. future auto speed control automation were independent variables. Perhaps the most interesting finding was the interaction between display and automation with respect to workload. When pilots were working with current automation, the position display was rated slightly higher on workload than the prediction display (though not statistically significant, $p=.054$). In the auto speed control automation condition, however, the prediction display produced higher workload than the position display. The researchers posited that because the prediction display resembled the 'green arc' used for altitude prediction, pilots were able to use it relatively easily when combined with familiar automation. Its advantage disappeared when the display was combined with new automation.

Designers should take these results to heart, as they highlight the need to avoid display cues that evoke newly inappropriate responses. They also suggest that pilots will need to be very careful when applying old mental models or 'how to' knowledge to new displays and automation. New expertise requires coherent processing - slightly different data displays must be examined thoroughly to ensure accurate comprehension, particularly when displays are coupled with new systems and procedures. A potential hazard is that pilots may think they 'recognize' and can use a display in familiar ways because it is similar to a current one (e.g., the prediction display is similar to the green arc), and may not be cognizant of differences in interpretation or function.

Caveats for expertise in automated environments. It is important to note that expertise does not offer the same advantages (i.e., the ability to use pattern-matching short-cuts) in the electronic world as it does in the naturalistic world. Experience may offer hints as to where to look for anomalies in technological environments, but it does not insulate domain experts from the need to navigate their way through layers of data when anomalies occur. This represents a shift to System2, or knowledge-based mode of problem solving. Performance and mistakes in knowledge-based cognition by experts resemble those of novices (Klein, et al., 2007; Reason, 1990).

Experience may even be counterproductive if accompanied by the tendency to rely on electronic data and systems in an intuitive manner. Riley (1994), for example, found that experienced pilots were *more* likely than students to inappropriately rely on automated systems after the systems had failed. More experience was correlated with a greater tendency toward automation bias in professional pilots (Mosier, Skitka, Heers, & Burdick, 1998). Experience can also work against a pilot, in that it may induce a false sense of coherence, or the tendency to see what one expects to see rather than what is there, as illustrated by the "phantom memory" phenomenon found in part-task simulation data (Mosier et al., 1998; Mosier, Skitka, Dunbar, & McDonnell, 2001).

System1 or System2? A critical component of expertise in the NextGen cockpit will be knowing when the situation is amenable to pattern recognition, and when it will require analysis. Pattern matching does not work when data are in digital format, are hidden below surface displays, or mean different things in different modes. Some current displays, however, may in fact be leading operators astray by fostering the assumption that data can be managed in a non-analytical fashion. This is an especially potent trap for highly experienced decision makers, who may be particularly vulnerable to errors if they 'see' patterns before they have sought out and incorporated all relevant information into their assessment. Training, then should emphasize the metacognitive monitoring of one's judgment and decision processes, to ensure that they are appropriate to the situation at hand.

Capitalizing on expert metacognition. The development of expertise for NextGen, as with any new task or system, will require a qualitative shift in processing and organization of knowledge (Tsang, 2003), and a continued emphasis on coherence skills. One aspect of expertise that professional pilots will take with them to the NextGen environment is metacognitive capabilities that surpass those of novices (Druckman & Bjork, 1991). Experts are aware of what they do and don't know, they plan ahead, they can manage their time and attentional resources, and they are able to monitor and revise their own problem-solving efforts (Cohen, 2011; Glaser, 1987; Tsang, 2003). Taking advantage of these abilities, training could target expanding these metacognitive skills for the new context, including training and practice in choosing the appropriate judgment strategy and in modifying it in response to situation events (Cannon-Bowers & Bell, 1997). In particular, pilots will need to monitor their cognitive processes to ensure they do not fall prey to System1 errors or biases.

Mitigating Automation Bias and Other Decision Biases

Debiasing Decision Making

Although the experience base of professional aircrews provides a layer of insulation against decision biases, pilots may still be susceptible to the threats inherent in cognitive shortcuts. For flight crews making decisions, biases such as representativeness, availability, confirmation bias, overconfidence, and automation bias are threats to front-end processes. Each of these may hinder accurate situation assessment, as they may impact the type and amount of information attended to or overlooked, as well as the weighting, categorization, and interpretation of that information. Confirmation bias, overconfidence, and automation bias may be particularly hazardous, as they threaten not only front-end situation assessment but also back-end processes such as making choices or mentally simulating outcomes of a course of action.

Researchers seeking ways to help individuals avoid decision biases have referred to these techniques as *debiasing* (Fischhoff, 1982; Hirt, Kardes & Markman, 2004; Sanna & Schwarz, 2003). Techniques have included increasing bias awareness, warning decision makers about relevant biases, providing feedback on decision making, improving decision training, "unfreezing" decision processes and changing them, and increasing incentives or motivation to strive for rationality in judgment and decision making. All of these techniques proved to have limited effectiveness, and their impact was found to vary by bias (e.g., Bazerman, 2002; Evans, 1989; Fischhoff, 1982; Humble, Keim, & Hershauer, 1992; Nisbett, Krantz, Jepson, & Ziva, 1983; Zakay, 1992). A short list of effective measures to eliminate biases has not been identified (Rehak, Adams, & Belanger, 2010). Instead, techniques were most successful if they were tailored to address specific biases, or as, Rehak, Adams, and Belanger (2010) suggested, if they were mapped onto the corresponding component or phase of decision making.

Research indicates that merely warning people about the influence of biases is not sufficient to eradicate them (Block & Harper, 1991; Fischhoff, 1982). On the other hand, training that provides individuals with feedback on their decisions and decision process has yielded some success (Alpert & Raiffa, 1982; Fischhoff, 1982; Sharp, Cutler, & Penrod, 1988). Bazerman (2002), for example, draws from the Lewin-Shein model of social change (Lewin, 1947; Schein, 1962), and discusses debiasing training in terms of unfreezing of old cognitive habits. Decision makers are provided with evidence of predictable errors in their judgment, receive instruction about biases and are given practice in new decision behaviors to integrate and refreeze new decision processes (Arnott, 2002)

Inducing System 2 processes to counteract biases. Because decision biases are the result of heuristics, or short-cuts in front- or back-end decision processes, successful debiasing interventions typically focus on eliciting or training more effortful cognitive (System 2) strategies (Milkman, Chugh, & Bazerman, 2009), for instance, by encouraging decision makers to consider alternatives or by holding them accountable for their decisions. Kahneman, Lovello, and Sibony (2011) proposed a 12-question process review for decision makers to forestall biases. In particular several items on their decision "quality control" checklist seem most useful to supporting analytical thinking by aircrews in NextGen operations:

⇒ *Check for saliency bias*. Is situation assessment/diagnosis overly influenced by salient cues, experiences or analogies? Is too much weight being put on salient indicators at the expense of less prominent ones? Are particularly salient successful/unsuccessful experiences really relevant and appropriate to the current situation?

⇒ *Check for confirmation bias*. Have credible alternatives to situation assessment been considered? Has disconfirming information been sought?

⇒ *Check for availability bias*. Is there any more information to be considered? If you had a checklist of information needed, what would be on it?

⇒ *Check for anchoring bias*. What was the initial diagnosis/assessment? Have later data and information been considered adequately? Have initial hypotheses been adjusted sufficiently?

Kahneman et al. recommend that because people usually do not recognize their own biases, the review should be done by individuals who are not part of the decision process. While this is not a practical recommendation for aircrews in the operational environment, it does suggest that the process would be effective only if it involves more than the deciding crewmember – so, for example, pilots could be trained to cross-check each others' judgment processes using these 'checks' as part of monitoring/challenging procedures.

Considering alternatives. This debiasing strategy requires individuals to generate alternative outcomes to scenarios (Lilienfeld, Ammirati, & Landfield, 2009). Instructing individuals to consider alternatives is believed to change their cognitive processing from automatic (heuristic) to a "rule-governed, controlled style of thinking" (Lilienfeld, Ammirati, & Landfield, 2009, p.393),. This technique was found to improve individuals' decision making beyond the original study domains or contexts (Galinsky & Moskowitz, 2000). For instance, Cohen et al. (1997) successfully used a variant of it to train military commanders' situation awareness.

Sanna, Schwarz and Stocker (2002) examined whether the "considering alternatives" technique could eliminate the hindsight bias. Study participants were asked to generate either 2 (easy condition) or 10 (difficult condition) alternatives to the outcome of a scenario, and to rate how likely these alternative were to occur. Hindsight bias was reduced when it was easy for participants to think of alternatives; however, the bias was exacerbated when alternatives were difficult to generate. Apparently participants equated the cognitive effort involved in the task with the inevitability of a given outcome.

Contradictory thinking. Given the plethora of data that will be available in the NextGen cockpit, pilots may limit information search to items that confirm their situation assessment. Pilots may also be more susceptible to confirmation bias if time-pressure is high, and/or if they have strong expectations, for instance, concerning their next clearance (e.g., Hooey & Foyle, 2001). Contradictory thinking may be an effective technique to mitigate confirmation bias, most notably to encourage decision makers to search for inconsistent information or to consider a wider set of decision scenarios (Fischhoff, 1982 in Arnott, 2006). Research by Burke (2006) on legal decision making suggests that lawyers could neutralize confirmation bias by acting as their own Devil's Advocates, reviewing cases from the perspective of the other side and creating counterarguments to their own position.

However, few researchers have examined how best to mitigate confirmation bias in a complex setting. Exceptions include work by Lehner, Adelman, Cheikes, & Brown (2008), who investigated the bias in a complex intelligence analysis task. They introduced a debiasing technique called Analysis of Competing Hypotheses (ACH), which required participants to list hypotheses and supporting evidence in a matrix. They found that ACH did reduce confirmation bias – but only in participants who did not have prior intelligence analysis experience. Consistent with earlier evidence in a complex tactical decision task (Tolcott, Marvin, & Lehner, 1989), Lehner et al. observed that the bias affected the weights that were given to confirming vs. disconfirming evidence, such that new evidence that supported participants' current beliefs was weighted more heavily than evidence that did not. Similarly, in a study simulating intelligence analysis, Wickens and colleagues (Wickens, Ketels, Healy, Buck-Gengler, & Bourne, 2010) found that warning participants about bias and emphasizing the age of intelligence information helped them to weigh the most recent cues more optimally; however, the instructions did not eliminate the anchoring heuristic. Anchoring is a ‘close cousin’ to confirmation bias and refers to the tendency of individuals to weigh more heavily cues and information that was received early compared to those obtained later.

Mental simulation. Mental simulation may be an especially useful mechanism to induce individuals to consider alternative outcomes. It is a strategy employed by expert decision makers and is amenable to training interventions. Galinsky and Moskowitz (2000) carried out a series of experiments in which study participants were asked to imagine ways an outcome could have turned out differently. The researchers hypothesized that counterfactual primes, or scenarios in which the outcome might have been different, would lead to mental simulations making relevant alternatives more accessible. In addition, it was hypothesized that the impact of counterfactual reasoning would generalize to other contexts. Participants were presented with counterfactual primes: at a rock concert a woman gives up her seat to get a better view of the stage; subsequently a surprise lottery is announced and the number of her new (or past) seat is the winning number. Participants in the control group only read about a woman who won (or did not win) a surprise lottery at a rock concert. Compared to control subjects, participants who had received a counterfactual prime were more successful and faster at solving a task (‘Duncker candle task’) whose solution depended on the non-typical use of objects. This finding suggests that the counterfactual scenario primed participants in the experimental condition to imagine alternative uses of objects.

While mental simulation is a strategy that expert decision makers employ in real world settings, they typically do not consider several alternative courses of action (Klein, 1989; 1993). Instead, they quickly size up a situation and generate an action plan, which they “run through” in their head. Only if this simulation leads to a problem or violates their expectations, do they consider alternative scenarios.

In the highly automated cockpit, decision-relevant patterns are not always easy to detect, even for experts, because of the nature of electronic data and information. Data may be distributed across several displays, and the meaning of data varies, for instance dependent on the system mode. Accordingly, pilots need to examine carefully all relevant data in order to avoid confirmation bias and accurately diagnose the situation. Errors related to confirmation bias were observed in aviation accidents and full-mission simulations using current automation (see Billings, 1996; Mosier et al., 1998; 2000). NextGen operational conditions, including increased reliance on electronic systems and time pressure, are likely to foster these types of errors. Techniques that promote the consideration of alternatives, such as mental simulation, Kahneman’s et al. process review questions, or Cohen’s et al. meta-recognitional skills training, may be useful countermeasures to guard against confirmation bias in complex environments such as NextGen operations.

Accountability demands

Accountability refers to making individuals responsible for providing justifications for their decision processes and choices (Tetlock & Kim, 1987; Skitka, Mosier & Burdick, 2000). Imposing accountability demands was found to sensitize decision-makers to the need to construct compelling justifications for their decision processes and choices. Social accountability can successfully ameliorate a broad array of cognitive biases, including primacy effects (Tetlock, 1983), the fundamental attribution error (Tetlock, 1985*a*), overconfidence effects (Tetlock & Kim, 1987), as well as the sunk cost effect (Simonson & Nye, 1992). Pre-decisional imposition of accountability demands cause decision makers to employ more multidimensional, self-critical, and vigilant information seeking, as well as more complex data processing, and has been shown to reduce cognitive "freezing" or premature closure on judgmental problems (Kruglanski & Freund, 1983), and to lead decision makers to employ more consistent patterns of cue utilization (Hagafors & Brehmer, 1983).

Mosier and colleagues (Mosier, Skitka, Heers & Burdick, 1998; Skitka, Mosier & Burdick, 2000) investigated whether accountability could ameliorate automation bias, the tendency to use automation as a heuristic replacement for vigilant information search, in students and in experienced airline pilots. It was hypothesized that increasing accountability for decisions and actions involving automation would increase vigilance in information processing strategies and cause operators to put more effort into verifying information and identifying appropriate responses than when accountability demands were not present.

A study with students (Skitka, Mosier & Burdick, 2000) provided evidence for the effectiveness of accountability as a mitigator of automation bias. Non-experienced participants using an automated aid completed computerized flight simulations that mimicked the monitoring and tracking tasks performed by professional pilots. Participants were made to feel accountable for speed, accuracy and overall performance, or were not made to feel accountable. Researchers found that being held accountable for overall performance or for accuracy led to fewer omission and commission errors and promoted increased vigilance, and that the lower rate of decision errors correlated with an increased rate of verifying information. This is significant because it suggests that accountability promoted more vigilant and coherent decision processes, which in turn resulted in more accurate decisions. Furthermore, there were no differences in response times, indicating that it did not take significant time to verify the information and avoid decision errors.

Accountability was experimentally manipulated in a part-task simulator study using experienced airline pilots (Mosier et al., 1998). Accountable participants were told that the purpose of the study was to analyze the *performance* of the pilot in heavily automated cockpits, that experimenters would be monitoring and evaluating their performance with respect to the use of automated flight systems, would be collecting performance data, and would ask them to justify their performance and strategies in the use of the automated systems in an interview following the experiment. Participants in the Non-Accountable condition were told only that the purpose of the study was to analyze the *role* of the pilot in heavily automated cockpits. To reinforce non-accountability, these pilots were also advised that, due to a breakdown of the data collection computer, no performance data would be collected. No mention of a formal interview or justification of performance was made. Monitors at the experimenter observation station were turned off while participants were in the room, and the video camera was pointed away from the subjects and was not turned on.

Researchers found partial support for the effectiveness of accountability as a debiasing technique for professional pilots. Although experimentally manipulated accountability demands did not significantly impact performance in this study, those professional pilots who reported on a debriefing questionnaire a higher internalized perception of 'accountability' for their performance and strategies of interaction with the automation were significantly more likely to double-check automated functioning against other cues, and committed fewer omission errors than those with a lower internalized perception.

Taken together, the studies indicate that interventions that emphasize accountability for decision processes may help to mitigate automation bias. Accountability accomplishes this by fostering more vigilant and thorough information search and integration, which will be critical in NextGen operations. The Mosier et al. (1998) study further suggests that a key factor for the success of accountability demands in professional pilots is that they be internalized rather than imposed. To accomplish this, airline cultures will need (a) to support the notion of accountability for decision processes, and also (b) to provide sufficient training on automation functioning so that pilots know what indicators should be checked, what information is relevant, and where to find information to verify or disconfirm their hypotheses.

Team Training – The Role of the Crew in NextGen Decision Making

The discussion to this point has dealt with decision making by individuals, focusing on individual biases and associated techniques for overcoming biased judgment and decision making. But in fact, aviation decisions are made in a crew context, primarily among the pilots on the flight deck, but also in larger teams involving the cabin crew, air traffic control, dispatch, maintenance, and – eventually – with other aircraft in a Next Gen environment.

Simply being in a crew or team context may reduce judgment or decision biases or errors because of the increased body of specialized knowledge and cognitive resources available. Team members can catch and correct each other's errors, contribute to more thorough information searches, evaluate evidence, and decide on a well-grounded course of action. But the social processes inherent in team functioning create their own costs and introduce team biases over and above those found in individual reasoning.

These team-level biases and techniques that have been developed for assuring effective team decision making by focusing specifically on team-level processes will be presented in a forthcoming chapter.

Training for Vigilance in Highly Automated Operations

Maintaining vigilance in NextGen operations will clearly be an issue. Flying 4D trajectories will require constant monitoring of automated systems, with few requirements for intervention in normal operations. Even the requirements for communication with ATC will be reduced, giving pilots few scheduled opportunities to check system and aircraft status. Many studies of vigilance have established the fact that the ability to monitor for the occurrence of infrequent, unpredictable events – such as automation errors - typically declines over time. Additionally, automated systems, especially those that are perceived to be highly reliable, may induce complacency in the operator, exacerbating the tendency toward non-vigilant monitoring (e.g., Parasuraman, Molloy, & Singh, 1993).

Training that includes automation failures has been proposed as a method to increase vigilance. Research participants who experienced an automation failure during task training exhibited less complacency than those who were simply told that the automation was not perfectly reliable (Bahner, Huper, & Manzey, 2008; Gore, et al., 2010). No magic bullet exists, however, to ensure continued vigilance across many hours of flying. The best avenue for promoting vigilance among pilots may well be through team training and crew resource management practices and procedures, as discussed below.

Meeting the Demands of a Highly Automated Cockpit While Maintaining Flying Skills

Increased flight deck automation may degrade pilots' flying skills and their ability to recover from automation failures. This concern was voiced by some participants in the FAA survey on aviation safety in NextGen (2010) who "urged a back-to-basics focus on stick-and-rudder skills…" (FAA, 2010, pp. 22-23). Reduced average pilot experience, reduced depth of piloting experience, and improved technology are all potential threats in the airline system (Bent & Chan, 2010). Moreover, it is possible for today's pilots to fly advanced aircraft for most of their careers, and never experience the need to acquire finely-tuned manual skills (Billings, 1996). Technology offers protection and increased efficiency for system operations, but reduces pilots' opportunities for on-the-job challenges and skill acquisition, highlighting the importance of simulator training that involves new classes of system failures.

A logical place to emphasize flying skills is during recurrent line-oriented flight training (LOFT). In particular, pilots need practice taking over when highly automated modes of operation fail, especially as the level of automation increases and requirements for pilot intervention during normal operations decrease (Endsley & Kaber, 1999). In addition to standard system failures, simulator training scenarios could include automation failures and events that require reversion to some level of manual flight. Additionally, during routine flights pilots can be encouraged to shut off the automation on a regular basis to enable practice on flying skills. These are potentially controversial suggestions, as NextGen operations depend on highly sophisticated automated systems, and airlines and aircraft manufacturers may prefer that pilots use full automation for reliability and efficiency; however, these suggestions are consistent with at least some airline statements of automation philosophy (Billings, 1996; Byrnes & Black, 1993).

Section 2. Automation Design for Decision Support: Current Technology and NextGen Recommendations

Automation Design for Decision Support

The implementation of new automated decision support systems in NextGen presents challenges to system designers, to expert users, and to the research community. The requirement for designers is to create systems that are 'expert-centered,' thereby facilitating and enhancing human expertise, particularly with respect to coherent use of automation. Expert operators must develop techniques to use automated aids efficiently and effectively, neither over- nor underestimating their capabilities. The challenge to the research community is to enable expert users and system designers to meet their respective goals, through empirical exploration and the establishment of principles for effective interaction between expert decision makers and automated systems. A body of literature exists in the area of engineering design to aid decision making—including several comprehensive chapters in *Reviews of Human Factors* volumes (Laughery & Wogalter, 2006; Rogers, Strong, & Fisk, 2006; Sheridan & Parasuraman, 2006; Smith, Bennett, & Stone, 2006). One theme that is consistent in the literature is that technological innovations to enhance decision making must be based on correct assumptions about human judgment processes and about how people think and use information (e.g., Maule, 2009).

Designing for Coherence

Presentation of Data and Information

Attention guidance. Implicit in the use of all relevant cues is an understanding of how the cues combine to create present awareness and prediction of a future situation. Attention distribution based on tactical considerations has been found to be systematically related to decision performance (Endsley & Smith, 1996). Moreover, attention to relevant cues is a predictor of both coherent diagnosis and accurate decisions. Consistent with this assessment is the finding in the medical field that the most effective decision support systems guided attention via clinical reminders (Garg et al., 2005).

Wickens (2003) discusses seven critical principles of display design related to information processing: information need, legibility, display integration/ proximity compatibility principle, pictorial realism, moving part, predictive aiding, and discriminability of status vs. command. These principles acknowledge the limitations in human attention and define characteristics that will facilitate processing of essential information. For example, the attitude indicator integrates attributes of pitch and roll (display integration/ proximity compatibility principle) so that the pilot can process both from the slant and vertical position of a single indicator (Vidulich, Wickens, Tsang, & Flach, 2010). Additionally, as pilots gain experience with cockpit displays, they develop selective attention strategies, directing their focus toward display areas that are relevant to the task at hand and where they expect to see changes (Senders, 1983).

Studies in diverse fields have demonstrated that experienced decision makers are likely to spend more time seeking out and attending to relevant cues than those who are less experienced (e.g., Cesna & Mosier, 2005; Khoo & Mosier, 2009; Salterio, 1996; Schriver, Morrow, Wickens, & Talleur, 2008). Experts also tend to "chunk" cues and to create patterns of relevant or structurally related cues (Chase & Simon, 1973; Endsley & Smith, 1996). Because attention is the critical first step in the front end of decision making, researchers and designers share the goals of supporting diagnosis and decisions through attention guidance (e.g., Horrey et al., 2006) or the identification of and training for efficient and coherent attention patterns (e.g., Schriver et al., 2008). This is critical to avoid decision process limitations and flaws such as attentional tunneling (the allocation of attention to a particular channel of information at the cost of other sources of information; e.g., Wickens & Alexander, 2009), "cognitive lockup" (the tendency to stay focused on one task even if a second task has higher priority; e.g., Kerstholt, Passenier, Houttuin, & Schuffel, 1996), attentional failures (e.g., Caird, Edwards, Creaser, & Horrey, 2005), or anchoring on early information with insufficient adjustment as new information comes in (Wickens, Ketels, Healy, Buck-Gengler, & Bourne, 2010). In NextGen, appropriate design of technological aids will enable coherent decision making by highlighting relevant data and information, helping pilots determine whether all information is consistent with a particular diagnosis or judgment, and providing assistance in accounting for missing or contradictory information.

Designers of display and notification systems must balance the need to gain the user's attention, generate rapid comprehension of the message, and encourage the most appropriate response (McCrickard, Catrambone, Chewar, & Stasko, 2003). It is commonly necessary to balance these objectives while also being sensitive to the context in which information is to be received. That context includes the physical environment, the technology used to enact the notification, and the other tasks currently occupying the operator. Furthermore, the techniques employed in delivering notifications influence the recipient's response. For instance, factors such as the form (e.g., numerical, graphical, narrative), ordering, context (e.g., positive vs. negative frame), and organization (e.g., in lists, tables, charts), as well as the vividness (i.e., the presence or absence of other salient visual cues, such as color and size of visual elements) of the information display will impact the attention to as well as comprehension of information.

Although current decision support systems do a reasonable job of guiding pilots' attention to system failures and potentially catastrophic anomalies (such as another aircraft in their path), they do not always facilitate attention to less salient information needed to verify aircraft mode or alerts and anomalies. As automated operations increase, NextGen pilots may be encouraged to rely uncritically on automated information, and may be susceptible to non-coherent or biased decision making. By carefully engineering the way in which data are presented to decision-makers for evaluation, it is possible to either overcome or exacerbate the use of heuristics and resultant biases, as well as other limitations that impact the quality of decisions that are made.

Designing to mitigate decision biases. Decision support systems (DSSs) can be designed in ways that mitigate decision biases, although some research has shown that DSSs do not always eliminate them (e.g., George, Duffy, & Ahuja, 2000). Arnott (2006), for example, proposes a DSS development cycle that uses cognitive bias theory as a focusing construct, and incorporates debiasing mechanisms similar to those discussed above (feedback, encouraging the search for disconfirming information, consideration of alternative scenarios or formulations of the problem) to overcome the negative effects of relevant biases. Interestingly, one step in Arnott's approach is to "reassure the user that the presence of biases is not a criticism of their cognitive abilities" (p. 65), a supportive notion that may make it more acceptable for operators to acknowledge susceptibility to decision biases and follow debiasing procedures.

Particular types of decision support interfaces may also ameliorate cognitive biases. Graphical interfaces rather than text, for example, have been shown to mitigate confirmation bias. In a war-games simulation, experienced Naval intelligence analysts were presented with combat scenarios either through an interactive graphical format or via text only. The scenarios presented to both groups contained a 'likely explanation' or hypothesis for the enemy activities detailed in the scenarios. Analysts were tasked with reviewing other intelligence concerning the enemy activities in order to determine how to proceed. Analysts that had been presented with the intelligence in graphical format were found to consider and incorporate more information that conflicted with the 'likely explanation' they had been provided than did their text-condition counterparts. The investigators posited that graphical presentation of the intelligence allowed the intelligence to remain more accessible and salient throughout the exercise, which in turn led to more disconfirming evidence being considered (Cook & Smallman, 2008).

DSSs can also be designed in ways that exacerbate cognitive biases. Systems that foster intuitive, System 1 thinking encourage heuristic processes and discourage analytical management. This is especially true when the interface does not provide the necessary information to establish or reinforce correct mental models of system functioning. Transparency will be an essential characteristic in the design of NextGen systems, as it is critical to the integration of relevant information and the management of systems.

Transparency. Many researchers have pointed out the need for more transparency and less hidden complexity in decision aiding systems (e.g., Woods, 1996; Sarter, Woods, & Billings, 1997). One of the biggest obstacles to human recognition of system errors is the poor quality of feedback about the activities of automated systems (Woods, 1994). For example, it is often difficult or impossible to trace the processes or predict the activities of automated flight systems. This is especially important when the logic of humans does not match the logic of the machine, a growing danger as computers and automation become more complex (Sheridan & Parasuraman, 2006). Principles of "human-centered automation" prescribe that the pilot must be actively involved, adequately informed, and able to monitor and predict the functioning of automated systems (Billings, 1996). To increase transparency, systems could provide roadmaps for locating relevant information and for tracking system recommendations, and point out sources of confirming or disconfirming information. If NextGen pilots are to rely on technological systems to support their decision making, they must be able to follow and understand the rationale underlying a DSS recommendation, find the information they need, integrate it appropriately into situation assessment and diagnoses, and predict the effect that any decision or action will have on aircraft systems and status.

Decision support for front-end versus back-end processes. Current and NextGen decision support systems may be geared toward either front-end situation assessment or back-end option selection. The literature suggests that technological decision aids will be most effective when they focus on enhancing situation assessment processes rather than merely prescribing decisions or actions. Work supporting the efficacy of support for front-end decision processes was provided in a study of decision aids for dealing with in-flight icing. Sarter and Schroeder (2001) found that both status and command displays led to improved performance over the baseline condition when accurate information was presented. However, when the decision aid provided inaccurate information, performance deteriorated below the baseline condition, and more so with the command display than with the status display. The authors proposed that the interaction could be related to the fact that status and command displays support different stages of decision-making. The status display supports the front end of the process, that is, diagnosis of a problem (situation awareness). In contrast, the command display supports the back end of the process, the choice or action selection stage. The authors posited that displaying back-end action suggestions encourages the pilot to cognitively remove him/herself from the diagnosis phase, and increases the risk that he or she may enter a “purely reactive mode” and follow system recommendations blindly (p. 581).

In other work, participants performing an air traffic control-related task exhibited superior performance when automated aids supported information acquisition and action implementation rather than performing cognitive functions such as information integration and analysis (Kaber, Perry, Segall, McClernon, & Prinzel, 2006). Back end support such as computer option generation can actually short-circuit front-end decision processes. Layton, Smith, and McCoy (1994), examined pilot use of a graphical flight planning tool, and found that computer generation of a suggestion or recommendation early in the course of problem evaluation significantly impacted decision processes and biased pilots towards the computer’s suggestion, even when the computer’s brittleness (e.g., in terms of an inadequate model of the “world”) resulted in a poor recommendation with potential adverse consequences. The use of automation for presentation of front-end information is critically important when fully reliable back-end decision automation cannot be guaranteed (Rovira, McGarry, & Parasuraman, 2007; see HART

Group, 2011, for more on automated decision aiding). If the pilot can achieve accurate situation assessment and awareness, resultant decisions are more likely to be appropriate.

Multiple modes. One prevalent design challenge for decision aid systems is avoiding overload in the visual channel, the mode most often used for information, warnings, and system displays. Sklar and Sarter (1999), for example, found that tactile cues, alone and redundant with visual cues, increased detection rates and response times to uncommanded mode transitions over visual cues alone. Interest in multimodal systems has increased, and these offer the potential for presenting data without data overload. Vibrotactile cues are being investigated as a way of providing directional cues in the vertical plane (Salzer, Oron-Gilad, Ronen, & Parmet, 2011). Issues in the use of multimodal displays include modality selection, mapping of modalities to tasks and information, combining, synchronizing, and integrating modalities, and adapting information to a multimodal presentation (Sarter, 2006). This intriguing design concept could provide high benefits in NextGen decision-making, as the proliferation of displays threatens to overwhelm visual processing capabilities, but further research is required to realize its potential.

Operator involvement. A major concern for NextGen is that highly automated systems that require little more than system monitoring will push pilots out-of-the-loop, hinder SA, and decrease pilots' ability to intervene when needed. Creating decision support systems that require operator involvement is one proposed remedy to this problem (Endsley & Kiris, 1995). Endsley and Kaber (1999), for example, found that people benefited most from a DSS that combined human generation of options and automated implementation of tasks. The need for pilot involvement in aircraft operation was emphasized as part of Billings' (1996) requirements for human-centered aircraft automation: "To command effectively, the human operator must be involved….To remain involved, the human operator must be appropriately informed…The human operator must be informed about automated systems' behavior" (p. 118-119).

Facilitating appropriate cognitive modes

One defining characteristic of NextGen cockpits will be the abundance of information available to the flight crew. Most of the new automation will focus on information presentation– pilots will have access to precise data concerning flight routes, traffic, or weather. A primary goal for designers of decision-support instrumentation and display systems should be to present and arrange information so that important data elements are made salient and accessible to decision-makers, contributing to development of their situation awareness (Ntuen, Park, & Gwang-Myung, 2010). One way to achieve this is through the use of data visualization tools, which present numerical data in a visually relevant information-congruent manner. The presentation of information is congruent when there is agreement between the meaning or significance of the information being conveyed and its mode of presentation (e.g., distance to target presented graphically is more congruent than if the same information were presented textually). For example, when experienced military personnel were provided information about enemy troop movements by means of interactive, electronic maps and diagrams that used animation to indicate combatant movement through terrain, they devised significantly better plans and did so in less time than did their counterparts who received the same amount of information via narrative, textual presentation (Ntuen et al., 2010).

The formats of automated displays will have predictable effects on the types of cognitive processes that will be elicited. For example, in a study of expert highway engineers, Hammond et al. (1997) found that specific characteristics of a task (i.e., display format and hidden relationships among variables within the task) tended to induce corresponding modes of responding: Judgments of highway aesthetics from filmstrips representing segments of the highways tended to be performed intuitively. Judgments of highway capacity via mathematical formulas were accomplished analytically. Quasi-rational cognition was induced by requiring judgments of highway safety from bar graphs. Because human response to information depends on how it is conveyed, it will be critical in NextGen to present information in a way that facilitates appropriate coherent decision processes. Designers of complex cockpit decision aids have taken advantage of visualization effects, presenting data in pictorial 'intuitive' formats whenever possible. Data are pre-processed and presented in a format that allows, for the most part, a wholistic view of aircraft positional and system states. Often, pictorial representations exploit human intuitive pattern-matching abilities, and allow quick detection of some out-of-parameter system states. This design philosophy seems to be consistent with the goals of workload reduction and information consolidation, and may facilitate faster and more accurate situation assessment and decision making in some situations.

However, other display requirements in the automated cockpit - making sense of electronic information systems, searching for hidden information, comprehending digital data, maintaining coordination between system states, and planning and programming future states - demand more analytical, coherence-based processing. Importantly, then, data that cannot be used intuitively (e.g., data that requires interpretation or takes on different meanings in different modes) should be presented in a format that facilitates analysis.

Numbers, modes, indicators, and computer readouts are some of the types of data that must be checked, processed, and integrated. Shortcuts may short-circuit the situation assessment process. Presenting data graphically when analysis is required may not elicit the appropriate response mode. For instance, when numerical data are plotted on a two-dimensional graph (consisting of two axes, X and Y), interpretation of the data may be manipulated by compressing or extending either of the axes, so that the physical placement of a number does or does not reflect its significance. For example, it may be that, while scores from 0 to 100 are possible for both attributes, a value of 70 on one scale may signal danger, while 70 on the other signals normal operation. Design of the layout of the scale on the screen may either help convey or mask this critical difference: for example, designers could avoid this problem by indicating 'normal' ranges through markers such as color bands on both dimensions. Sun, Li, & Bonini found that inappropriate cognitive modes can be eliminated or reduced by requiring the evaluators of the information to justify their judgments in explicit terms, thus prompting them to switch to a more analytical mode of processing (Sun, Li, & Bonini, 2010). Interestingly, this relates to and provides additional support for the notion that accountability for decision processes, as discussed above, may mediate pilot interaction with automated displays and mitigate use of the automation heuristic and resultant automation bias.

Supporting Metacognition

In addition to presenting data and information appropriately, decision-aiding technology should be designed to enhance individual and crew metacognition—that is, it should help people to be aware of how they are thinking and are making judgments, and whether or not their process is appropriate, coherent, and accurate. Decision aids need to be adaptive not only in the *when* of aiding, but also in the *how*. Perhaps one of (if not *the*) most important functions of decision aids may be to help operators to monitor their decision making by prompting processes that guard against human error or bias (e.g., *check system parameters, look for disconfirming evidence,* *what are the consequences of this action, do you really want to…*) as well as by providing information in an appropriate manner. Guerlain et al. (1999), for example, proposed an interactive critiquing system with which the human decision maker can submit his/her diagnosis to the automation to be checked. Also, in order to be truly effective, decision aids must help operators to navigate within and between decision contexts, especially when they are initiating a new task or changing contexts (e.g., when pilots transition from a visual to an instrument approach or when shifting from monitoring to active control), and to adapt their own cognition effectively to changing demands of the situation.

Facilitating teamwork and shared mental models. The increased emphasis on teams in NextGen – both co-located and distributed - has created additional demands for decision aids. In particular, when automated aids can function under many different modes and aircraft status can vary depending on weather, traffic, and operational priorities, it is critical that all team members share the same correct model of current states. The implications of mode or status changes by one team member must be recognized by all team members, and everyone must know when actions need to be or have been taken (Moray, 1992; Sheridan & Parasuraman, 2006). This will be increasingly important in NextGen as roles and responsibilities shift among team members. Decision aids, then, must facilitate team cognition by reminding all users of current state and what they can expect to happen in future. Failure to do this can lead to inaccurate mental models and a breakdown of team coordination and decision making. This topic will be further elaborated in the forthcoming chapter on team-level decision making.

Avoiding the De-Skilling of Decision Making Skills

NextGen operations pose a threat not only to the flying skills of professional pilots, but also to their decision making skills. Issues around the potential erosion of decision-making skills are certain to take on more importance and relevance as NextGen operations evolve. A long-range hazard of extensive use of automated decision-aiding systems is what Hart (1992) has termed the "deskilling" of experts. As NextGen automation becomes more prevalent and more autonomous, deskilled experts may no longer have the degree of expertise to judge when the system is performing correctly, to carry out the necessary degree of supervision over decision functions, or to take over from automation when necessary (Morgan, 1992).

Historically, automated decision support systems have functioned as *consultants* in the decision-making process. Through tradition or regulation, the user has maintained ultimate decision-making authority and borne responsibility for the outcomes of decisions. As these systems become more capable of autonomous functioning, and are given more power in the guidance of human activity, however, the question of who or what should have responsibility for decision making becomes more complicated. The ubiquity of automated support systems may subtly erode the decision maker's role, and foster almost an abdication of decision responsibility to the systems. Human operators may see reliance on DSSs as the most efficient way to make decisions, particularly in times of high workload. In fact, because an explicit objective of automated systems is to reduce human error, automating a decision-making function may communicate to the operator that the automation *should* be given primary responsibility for decisions.

The issue of responsibility in the use of automation and decision aids is complex and at times confusing. Who--or what--should be liable if the human-automation system fails? Although responsibility for actions typically rests with the user (Mockler, 1989), the legal status of the user of automated system technology is not always clear (Will, 1991), especially in instances in which the user ignores or overrides the advice of the system. For "certain sensitive, delicate or hazardous tasks [such as aircraft requiring fast and accurate response beyond human capability], it may be unreasonable not to rely upon an expert system" (Gemignani, 1984, p. 1045). In fact, airline policies may encourage using the highest levels of automation in an effort to shift responsibility from pilots to systems (personal communication with C. Abbott, 2011). In the future, lawsuits may result if system aids are consulted and fail to perform correctly, or give inaccurate or misleading indications; conversely, operators may be liable for the nonuse of an available system or for ignoring its recommendations when they are correct (Zeide & Liebowitz, 1987).

Caveats and Conclusions

> *People seem to believe that greater investments in automation promise lower expenditures on developing human expertise. However, the data consistently show that the impact of new levels and types of automation is new knowledge requirements for people in the system as their role changes to more of a manager and anomaly handler (Sarter et al., 1997, p. 334)*

Throughout this report, we have emphasized that the key driver in NextGen decision making will be automation. Trajectory-based operations (TBO), for example, will require 4-dimensional automated navigation capabilities. Information and decision support systems will provide data about aircraft state, will help pilots synthesize the data, and may offer recommendations for action. Automated sources such as SWIM (system-wide information management) will make up-to-the minute weather information easily accessible. Automated broadcasting systems (e.g., ADS-B, automatic dependent surveillance-broadcast) will facilitate accurate tracking of traffic. Automated navigational systems will enable flight paths that are precise with respect to both position and time. Sophisticated datalink systems will allow crews to input new clearances or flight changes with the push of a button. Automation will enable all-weather operations that have not been possible in the past, such as CSPR approaches and landings. In the NextGen operational environment, electronic data and information will supplant most out-the-window cues, and pilot attention will be much more inside the cockpit than ever before.

Although the benefits of NextGen automation are undeniable, some caveats must be acknowledged and addressed in crew training as well as system design.

'Mind the Gap' - Pilots Need Accurate Mental Models of Automated Systems

Current pilot automation training focuses on procedural knowledge rather than functional relationships. The 'cookbook,' step-by-step approach to managing automated systems results in large gaps in pilots' understanding of how systems work and limits their ability to cope with any situations outside of normal, standard operations (Sarter & Woods, 1994). More than a decade ago, Billings (1996) and others (Sarter & Woods, 1994) called for training on how systems work as well as how to work the systems, and the potential costs of ignoring their call in NextGen operations are enormous. Pilots will still form the last line of defense in the event of anomalies or emergencies, and they will need the resilience provided by a deep level of system knowledge during non-normal operations. This will be especially important as new responsibilities are shifted from ATC to flight crews.

Electronic Systems Require System 2 Processing

Electronic data are not amenable to System 1 intuitive processes. Pilots must be aware that even when data are presented in graphic format or seem to present patterns for recognition, their meaning must be interpreted analytically in the context of automation mode, flight phase, and system status. Moreover, expert pilots may not be able to rely on System 1 short-cuts such as pattern-matching or heuristics in NextGen because the architecture of the knowledge space will have changed. Training programs will need to focus on systems knowledge and coherence skills - seeking out relevant data and information and ensuring that they form a consistent 'story' of the situation - and system design must facilitate the front-end analytical processing required when dealing with electronic data and information.

Systems Awareness Is Key to Situation Awareness

Situation awareness in the NextGen cockpit cannot be achieved without systems awareness. Because most of the information for SA will reside within electronic systems, an awareness of what systems are doing and why, what mode they are in, and what they will do next is essential to accurate assessment and awareness. Training and design will need to support accurate system models and should emphasize systems awareness as well as the integration of data and information inside the cockpit with elements outside of the cockpit.

What's New and Different? Changes Must Be Highlighted

One potentially hazardous aspect of NextGen is that it is to some extent an updated paradigm, mixing current and new automation and displays. Changes to systems and procedures are being introduced incrementally, and much of the 'new' automation actually comprises enhancements to current automation. The FMS, for example, will be upgraded to a 4D management system, and will enable autoload of clearances. NextGen systems and displays will be somewhat different from current systems and displays, but not completely so, luring pilots into old patterns of thinking and interaction when these behaviors are no longer appropriate. Pilots with the most experience with old systems may be most susceptible to this trap, as their patterns will be more strongly entrenched and they will be susceptible to 'strong-but-wrong' erroneous behavior that is more consistent with past practice than current requirements (Reason, 1990). Training and design must include a focus on what is different in systems and displays. Pilots must be aware of nuances of new automation, differences in its functioning, and situations in which it is likely to mislead them.

'Monitor and Challenge' Also Applies to Automation

NextGen automation will provide flight crews with amazing tools for safe and efficient flight. With new tools, however, come new tasks and duties. For example, RNP (Required Navigational Performance) for TBO will confer increased autonomy on flight crews with respect to navigation, along with the concomitant responsibilities for flight monitoring and self-spacing. Perhaps the primary responsibility of the pilots will be to use the new tools wisely in their decision making processes. Crew resource management programs train pilots to monitor each other, and to challenge each other when anything seems awry. A similar approach to automation will increase the probability that the promise of increased safety and efficiency in NextGen operations will be realized.

ACKNOWLEDGEMENTS

Thanks to SFSU student assistants Kerry Cunningham, Alec Munc, and Christine Thomas.

REFERENCES

Alpert, M. and Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 294-305). Cambridge, UK: Cambridge University Press.

Arnott, D. (2006). Cognitive biases and decision support systems development: A design science approach. *Information Systems Journal, 16,* 55–78.

Bahner, E., Huper, A-D., & Manzey, D. (2008). Misuse of automated aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies, 66,* 688-699.

Ballard, D H., Hayhoe, M. & Pelz, J. (1995). Memory representations in natural tasks. *Journal of cognitive neuroscience, 7(1)*, 66-80.

Bayes, T. (1958). Essay towards solving a problem in the doctrine of chances. *Biometrika, 45*, 293-315. (Reprinted from *Philosophical Transactions of the Royal Society,* 1763*, 53,* 370-418.

Bazerman, M.H. (2002) *Judgment in Managerial Decision Making* (5th ed.). Wiley, New York.

Bent, J., & Chan, K. (2010). Flight training and simulation as safety generators. In E. Salas & D. Maurino (Eds.), *Human factors in aviation, 2nd Edition* (pp. 293-334). Burlington, MA: Academic Press (Elsevier).

Billings, C. E. (1996). *Human-centered aviation automation: Principles and guidelines*. (NASA Technical Memorandum #110381) NASA Ames Research Center, Moffett Field, CA.

Block, R. A., & Harper, D. R. (1991). Overconfidence in estimation: testing the anchoring-and-adjustment hypothesis. *Organizational Behavior and Human Decision Processes*, 49, 188-207.

Bolstad, C. A., Cuevas, H. M., Connors, E. S., Foltz, P. W., Gonzalez, C., Lau, N. K. C., & Warwick, W. J. (2010). Advances in Modeling Situation Awareness, Decision Making, and Performance in Complex Operational Environments. *Proceedings of Human Factors and Ergonomics Society 54th Annual Meeting* - 2010, 1047–1051.

Burke, A. (2006) Neutralizing cognitive bias: An invitation to prosecutors. *NYU Journal of Law & Liberty, 2,* 512-530.

Byrnes, R. E., & Black, R. (1993). Development and implementing CRM programs: The Delta experience. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego, CA: Academic Press.

Caird, J. K., Edwards, C. J., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors, 47(2)*, 235-249.

Cannon-Bowers, J. Bell, H. H. (1997). Training decision makers for complex environments: Implications of the Naturalistic Decision Making perspective. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 99-110). Mahwah, NJ: Lawrence Erlbaum Associates.

Cesna, M., & Mosier, K. L. (2005). Using a prediction paradigm to compare levels of expertise and decision making among critical care nurses. In B. Brehmer, R. Lipshitz, & H. Montgomery (Eds.), *How do professionals make decisions*? (pp. 107-118). NY: Lawrence Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.

Cohen, M. (2011). Knowns, known unknowns, and unknown unknowns: Synergies between intuitive and deliberative approaches to time, uncertainty, and information. In K. L. Mosier & U. M. Fischer (Eds.), *Informed by knowledge: Expert performance in complex situations*. NJ: Taylor & Francis.

Cohen, M. S., Freeman, J. R., & Thompson, B. B. (1997). Training the naturalistic decision maker. In C. E. Zsambok, & G. Klein (1997). *Naturalistic decision making* (pp. 257-268). Mahwah, NJ: Lawrence Erlbaum Associates.

Cook, M. B., & Smallman, H. S. (2008). Human factors of the confirmation bias in intelligence analysis: Decision support from graphical evidence landscapes. *Human Factors, 50(5)*, 745-754.

Druckman, D., & Bjork, R. A. (Eds.). (1991). *In the mind's eye: Enhancing human performance*. Washington, DC: National Academy.

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman, & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163-181). New Jersey: Lawrence Erlbaum Associates.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 3-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M.R. & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in dynamic control task. Ergonomics, 42, 462-492.

Endsley, M. R., & Kiris, E. O. (1995). Situation awareness global assessment technique (SAGAT) TRACON air traffic control version user's guide. Lubbock: Texas Tech University.

Endsley, M. R. & Smith, R. (1996.) Attention Distribution and Decision Making in Tactical Air Combat. Human Factors, 38(2), 232-249.

Evans, J.St.B.T. (1989) *Bias in Human Reasoning: Causes and Consequences.* London: Lawrence Erlbaum Associates.

FAA (2009). *NextGen Mid-Term Concept of Operations for the National Airspace System, Version 1.0*. Washington, DC: FAA Air Traffic Organization NextGen & Operations Planning, Research & Technology Development, Air Traffic Systems Concept Development.

FAA (2010). Answering the Call to Action on Airline Safety and Pilot Training.

Fischer, U., Orasanu, J., & Davison, J. (2003). Why do pilots take risks? Some insights from a think-aloud study. In M. J. Cook (Ed.), *Proceedings of the Human Factors of Decision Making in Complex Systems Meeting* (pp. 44-46). Dundee, Scotland: University of Abertay.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge, UK: Cambridge University Press.

Galinsky, A.D. & Moskowitz, G. B. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. Journal of Experimental Social Psychology, 36, 384-409.

Garg, A, X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., et al (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes. *Journal of the American Medical Association, 293,* 1233-1238.

Gemignani, M. (1984). Laying down the law to robots. *San Diego Law Review, 21,* 1045.

George, J. F., Duffy, K., & Ahuja, M. (2000). Countering the anchoring and adjustment bias with decision support systems. *Decision Support Systems*, 29, 195-206.

Gore, B. F., Hooey, B. L., Wickens, C. D., Sebok, A, et al. (2010). *Identification of Pilot Performance Parameters for Human Performance Models of Off-Nominal Events in the NextGen Environment* (NASA Technical Publication CR-2010-216411). Hanover, MD: NASA Center for AeroSpace Information

Glaser, R. (1987). Thoughts on expertise. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 81-94). Norwood, NJ: Ablex.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics.* Huntington, NY: Krieger.

Guerlain, S. A., Smith, P., Obradovich, J., Rudmann, S., Strohm, P., Smith, J. W. et al. (1999). Interactive critiquing as a form of decision support: An empirical evaluation. Human Factors, 41, 72-89.

Hagafors, R. & Brehmer, B. (1983). Does having to justify one's decisions change the nature of the decision process? *Organizational Behavior and Human Performance*, 31, 223-232.

Hamm, R. M. (1988). Moment-by-moment variation in expert's analytic and intuitive cognitive activity. *IEEE Transactions on Systems, Man, and Cybernetics, 18(5*), 757-776.

Hammond, K. R. (1993). Naturalistic decision making from a Brunswikian viewpoint: Its past, present, future. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods*. NJ: Ablex Publishing Co.

Hammond, K. R. (1996). *Human Judgment and Social Policy*. New York: Oxford Press.

Hammond, K. R. (1999). Coherence and correspondence theories of judgment and decision making. In T. Connolly, H. R. Arkes, & K. R. Hammond (Eds.), Judgment and decision making: An interdisciplinary reader (pp. 53-65). NY: Cambridge University Press.

Hammond, K. R. (2000). *Judgments under stress*. New York: Oxford Press.

Hammond, K. R. (2007). *Beyond rationality: The search for wisdom in a troubled time*. NY: Oxford University Press.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1997). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. In W. M. Goldstein & R. M. Hogarth (Eds.), *Research on judgment and decision making: Currents, connections, and controversies* (pp. 144-180). Cambridge: Cambridge University Press.

Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). *Human judgment and decision making: Theories, methods, and procedures*. NY: Praeger.

Hart, A. (1992). *Knowledge Acquisition for Expert Systems*. McGraw Hill, 2nd edition.

HART Group (2011). *Automation in the Cockpit: Toward a Human-Automation Relationship Taxonomy*. Interim Report, FAA Cooperative Agreement #DTFAWA-10-C-00084.

Hirt, E.R., Kardes, F.R. & Markman, K.D. (2004). Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. Journal of Experimental Social Psychology, 40, 374-383.

Hooey, B. L., & Foyle, D. C., (2001). A post-hoc analysis of navigation errors during surface operations: Identification of contributing factors and mitigating solutions. Proceedings of the Eleventh International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University.

Horrey, W. J., Wickens, C. D., Strauss, R., Kirlik, A., & Stewart, T. R. (2006). Supporting situation assessment through attention guidance and diagnostic aiding: The benefits and costs of display enhancement on judgment skill. In A. Kirlik (Ed.), *Adaptive perspectives on human-technology interaction* (pp. 55-70). New York, NY: Oxford University Press .

Humble, J.E., Keim, R.T. & Hershauer, J.C. (1992) Information systems design: An empirical study of feedback effects. *Behaviour and Information Technology, 11*, 237-244.

Kaber, D. B., Perry, C. M., Segall, N., McClernon, C. K., & Prinzel III, L. J. (2006). Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. *International Journal of Industrial Ergonomics, 36*, 447-462.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93*, 1449-1475.

Kahneman, D., Lovallo, D., & Sibony, O. (2011). Before You Make That Big Decision.. *Harvard Business Review*, 89(6), 50-60.

Kerstholt, J., Passenier, P., Houttuin, K. & Schuffel H. (1996.) The Effect of a Priori Probability and complexity on Decision Making in a Supervisory Control Task. Human Factors, 38(1), 65-78.

Khoo, Y-L., & Mosier, K. L. (2009). The impact of time pressure and experience on information search and decision making processes. *Journal of Cognitive Engineering and Decision Making*.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), Advances in Man-Machine System Research, Volume 5 (pp. 47-92). Greenwich, CT: JAI Press.

Klein, G. (1993). Recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood & C. Zsambok (Eds.), Decision making in action: Models and methods (pp. 103-147). Norwood, NJ: Ablex.

Klein, G., (2000). Decision skills training for the aviation community. *Proceedings of the Fifth Australian Aviation Psychology Symposium*, November, Manly, Australia.

Klein, G. A., Calderwood, R., & Clinton Cirocco, A. (1986). *Rapid decision making on the fire ground* (KA-TR-84-41-7). Yellow Springs, OH: Klein Associates Inc. (Prepared under contract MDA903-85-G-0099 for the U.S. Army Research Institute, Alexandria, VA)

Klein, G., Phillips, J.K., Rall, E., & Peluso, D.A. (2007). A data-frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise Out of Context: Proceedings of the 6th International Conference on Naturalistic Decision Making* (pp. 113-155). Mahwah, NJ: Lawrence Erlbaum Associates.

Kruglanski, A. W., & Frennd, T. (1983). The freezing and unfreezing of lay inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimenta1 Social Psychology*, 14, 448-468.

Laughery, K. R., & Wogalter, M. S. (2006). Designing Effective Warnings. In R. C. Williges (Ed.), *Reviews of human factors and ergonomics, Volume 2* (pp. 241-271). Santa Monica, CA: Human Factors and Ergonomics Society.

Layton, C., Smith, P. J., & McCoy, C. E. (1994). Design of a cooperative problem-solving system for en-route flight planning: An empirical evaluation. *Human Factors, 36*, 94-119.

Lehner, P. E., Adelman, L., Cheikes, B. A., & Brown, M. J. (2008). Confirmation bias in complex analyses. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 38*, 584-592.

Levin, D. & Simons, D J. (1997). Change blindness. *Trends in cognitive sciences, 1(7)*, 261-267.

Levin, D., Momen, N., Drivdahl, S., & Simons, D. (2000). Change blindness: The metacognitive error of overestimating change-detection ability. *Visual cognition, 7*, 397-412.

Lewin, K. (1947a). Frontiers in group dynamics: Concept, method, and reality in social science; social equilibria and social change. *Human Relations*, 1, 5–41.

Lilienfeld, S. O., Ammirati, R. & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science, 4*, 390-398.

Maule, J. (2009). Can computers help overcome limitations in human decision making? *Proceedings of NDM9, the 9th International Conference on Naturalistic Decision Making.* London, UK: British Computer Society.

McCrickard, D., Catrambone, R., Chewar, C., & Stasko, J. T. (2003). Establishing tradeoffs that leverage attention for utility: empirically evaluating information display in notification systems. *International Journal of Human-Computer Studies*, 58, 547–582.

Milkman, K.L., Chugh, D., Bazerman, M.H. (2009). How can decision making be improved? *Perspectives on Psychological Science*, *4*(4), 379-383.

Mockler, R. J. (1989). *Knowledge-based systems for management decisions*. Englewood Cliffs, NJ: Prentice-Hall.

Moray, N. (1992). Flexible interfaces can promote operator error. In J. Kragt (Ed.), *Case studies in ergonomics* (pp. 49-64). London: Taylor & Francis.

Morgan, T. (1992). Competence and responsibility in intelligent systems. *Artificial Intelligence Review, 6,* 217-226.

Mosier, K. L. (2002). Automation and cognition: Maintaining coherence in the electronic cockpit. In E. Salas (Ed.), *Advances in Human Performance and Cognitive Engineering Research, Volume 2.* Elsevier Science Ltd. (pp. 93-121)

Mosier, K. L. (2009). Searching for Coherence in a Correspondence World. *Judgment and Decision Making (4), 2*,

Mosier, K., Fischer, U., & Orasanu, J. (2011). *Flight Crew Decision Making: Now and NextGen.* Interim Report, FAA/NASA NextGen Flight Deck Human Factors Research Interagency Agreement #DTFAWA-10-X-80005, Annex 9.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. D. (1998). Automation bias: Decision making and performance in high-tech cockpits. *International Journal of Aviation Psychology, 8*, 47-63.

Mosier, K. L., Skitka, L. J., Dunbar, M., & McDonnell, L. (2001). Air crews and automation bias: The advantages of teamwork? *International Journal of Aviation Psychology, 11*, 1-14.

Murray, S. R. (1997). Deliberate decision making by aircraft pilots: A simple reminder to avoid decision making under panic. *International Journal of Aviation Psychology, 7*, 83-100.

Nisbett, R.E., D.H. Krantz, C. Jepson, and Z. Kunda (1983), The use of statistical heuristics in everyday inductive reasoning," *Psychological Review*, *90*(4), 339–63.

Ntuen, C., Park, E., & Gwang-Myung, K. (2010). Designing an Information Visualization Tool for Sensemaking. *International Journal of Human-Computer Interaction*, 26(2), 189–205.

Orasanu, J. (2010). Flight crew decision making. In B. G. Kanki, R. L. Helmreich, and J. Anca (Eds.), *Crew Resource Management, Second Edition* (pp. 147-179). San Diego: Academic Press/Elsevier.

Parasuraman, R. M., Molloy, R., & Singh I. L. (1993). Performance consequences of automation induced" complacency." *International Journal of Aviation Psychology, 3,* 1–23.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000, May). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, Man, & Cybernetics, 30*, 286-297.

Prince, C., & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 337-366). San Diego, CA: Academic Press.

Rehak, L. A., Adams, B., & Belanger, M. (2010). Mapping biases to the components of rationalistic and naturalistic decision making. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society (pp. 324–328).

Rensink, R A. (2002). Change detection. *Annual Review of Psychology, 53*(1), 245-77.

Riley, V. (1994). A theory of operator reliance on automation. In M. Mouloua & R. Parasuraman (Eds.), *Automation and human performance: Theory and applications* (pp. 19-35). New Jersey: Lawrence Erlbaum Associates.

Rogers, W. A., Stronge, A. J., & Fisk, A. D. (2006). Technology and Aging. . In R. S. Nickerson (Ed.), *Reviews of Human Factors and Ergonomics, Volume 1* (pp. 130-171). Santa Monica, CA: Human Factors and Ergonomics Society.

Rovira, E., McGarry, K., & Parasuraman, R. (2007). Effects of imperfect automation on decision making in a simulated command and control task. *Human Factors, 49*, 76-85.

Salas, E., Cannon-Bowers, J. A., Fiore, S. M., & Stout, R. J. (2001). Cue-recognition training to enhance team situation awareness. In M. McNeese, E. Salas, & M. Endsley, (Eds.), *New trends in collaborative activities: Understanding system dynamics in complex environments* (pp. 169-190). Santa Monica, CA: Human Factors and Ergonomics Society

Salterio, S. (1996.) Decision Support and Information Search in a complex Environment: Evidence from Archival Data in Auditing. *Human Factors*, 38(3), 495-505.

Salzer, Y., Oron-Gilad, T, Ronen, A, & Parmet, Y. (2011). Vibrotactile "on-thigh" alerting system in the cockpit. *Human Factors, 53,* 118-131.

Sampanes, A., Bridgeman, B. & Tseng, P. (2008). The role of gist in scene recognition. *Visions Research, 48(21),* 2275-2283

Sanna, L. J., & Schwarz, N. (2003). Debiasing the hindsight bias: The role of accessibility experiences and (mis)attributions. *Journal of Experimental Social Psychology*, 39, 287–295.

Sarter, N. (2006). Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics (36)*, 439-445.

Sarter, N., & Schroeder, B. (2001.) Supporting Decision Making and Action Selection under Time

Pressure and Uncertainty: The Case of In-Flight Icing. *Human Factors*, 43(4) 573-583.

Sarter, N., & 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*, 5-19.

Sarter, N B., Woods, D. D., & Billings, C. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of human factors/ergonomics* (2nd ed., pp. 1926-1943). New York: Wiley.

Schriver, A. T., Morrow, D. G., Wickens, C. D., Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50,* 864~878.

Schein, E.H. (1962) Management development as a system of influence. *Industrial Management Review,* 2, 59-77.

Schein, E. H. (1987) *The clinical perspective in fieldwork.* Sage Publications, Newbury Park, CA.

Senders, J. W. (1983). Models of visual scanning processes. *Bulletin of the Psychonomic Society, 21*, 340.

Sharp, G. L., Cutler, B. L., & Penrod, S. D. (1988). Performance feedback improves the resolution of confidence judgments. *Organizational Behavior and Human Performance,* 42, 271–283.

Sheridan, T. B., & Parasuraman, R. (2006). Human-automation interaction. In R. S. Nickerson (Ed.), *Reviews of Human Factors and Ergonomics, Volume 1* (pp. 89-129). Santa Monica, CA: Human Factors and Ergonomics Society.

Sherry, L., & Polson, P. G. (1999). Shared models of flight management system vertical guidance. *International Journal of Aviation Psychology, 9,* 139-153.

Simons, D J. (1996). In sight, out of mind: When object representations fail. *Psychological science, 7(5)*, 301-305.

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition, 7,* 1-15.

Simonson, I. & Nye, P. 1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes, 51,* 416-446.

Skitka, L. J., Mosier, K. L., Burdick, M., & Rosenblatt, B. (2000). Automation bias and errors: Are crews better than individuals? *International Journal of Aviation Psychology, 10(1),* 83-95.

Skitka, L. J., Mosier, K. L., Burdick, M., Rosenblatt, B. (2000). Automation bias and errors: Are crews better than individuals? *International Journal of Aviation Psychology, 10(1),* 83-95.

Sklar, A. E., & Sarter, N. B. (1999). Good Vibrations: Tactile Feedback in Support of Attention Allocation and Human-Automation Coordination in Event-Driven Domains. *Human Factors (41),* 543-552.

Smith, P. J., Bennett, K. B., & Stone, R. B. (2006). Representation aiding to support performance on problem-solving tasks. In R. C. Williges (Ed.), *Reviews of human factors and ergonomics, Volume 2* (pp. 74-108). Santa Monica, CA: Human Factors and Ergonomics Society.

St. John, M., Smallman, S., Manes, D., Feher, B., & Morrison, J. (2005) Heuristic automation for decluttering tactical displays. *Human Factors, 47*(3), 509.

St. John, M., Smallman, S., & Manes, D. (2005). Recovery from interruptions to a dynamic monitoring task: The beguiling utility of instant replay. Presented at the 49th Annual Meeting of the *Human Factors and Ergonomics Society*.

Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology*

Quarterly, 48, 227-236.

Tetlock, P. E. (1985). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly,* 48,74-83.

Tetlock, P. E., & Kim, .II.. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology, 52,* 700-709.

Tolcott, M.A., Marvin, F.F., & Lehner, P.E. (1989). Expert decision making in evolving situations. *IEEE Transactions on Systems, Man, and Cybernetics, 19(3)*, 606-615.

Treisman, M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97.

Tsang, P. S. (2003). Assessing cognitive aging in piloting. In P. S. Tsang & M. A. Vidulich (Eds.) *Principles and practice of aviation psychology* (pp. 507-546). Mahwah, NJ: Lawrence Erlbaum Associates.

Verma, S., Martin, L., Lozito, S., Kaneshige, J., Ballinger, D., Kozon, T., Cheng, L., Sharma, S., & Subramanian, S. (2011). *Integrated pilot and controller procedures: Aircraft pairing for simultaneous approaches to closely spaced parallel runways*. Presentation at the Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011).

Vidulich, M. A., Wickens, C. D., Tsang, P. S., & Flach, J. M. (2010). Information processing in aviation. In E. Salas & D. Maurino (Eds.), *Human factors in aviation, 2nd Edition* (pp. 175-216). Oxford, UK: Elsevier.

Wickens, C. D. (2003). Aviation displays. In P. S. Tsang & M. A. Vidulich (Eds.), *Principles and practice of aviation psychology* (pp. 147-200). Mahwah, NJ: Lawrence Erlbaum Associates.

Wickens, C.D. & Alexander, A, L. (2009). Attentional tunneling and task management in synthetic vision displays. *International Journal of Aviation Psychology*, 19(2), 1-17.

Wickens, C. D., Ketels, S. L., Healy, A. F., Buck-Gengler, C. J., & Bourne, L. E., Jr. (2010). The anchoring heuristic in intelligence integration: A bias in need of de-biasing. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Will, R. P. (1991), True and false dependence on technology: Evaluation with an expert system. *Computers in Human Behavior, 7*, 171-183.

Woods, D. D. (1996). Decomposing automation: Apparent simplicity, real complexity. In R. Parasuraman, & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 3-18). New Jersey: Lawrence Erlbaum Associates.

Woods, D. D., & Sarter, N. B. (2000). Learning from automation surprises and "going sour" accidents. In N. B. Sarter & R. Amalberti (Eds.), *Cognitive engineering in the aviation domain* (pp. 327-353). Mahwah, NJ: Lawrence Erlbaum Associates.

Zakay, D. (1992). The influence of computerized feedback on overconfidence in knowledge. *Behaviour and Information Technology*, 11, 329–333.

Zeide, J. S., & Liebowitz, J. (1987, Spring). Using expert systems: The legal perspective. *IEEE Expert*, p. 19-22.

Zsambok, C. E., & Klein, G. (1997). *Naturalistic decision making.* Mahwah, NJ: Lawrence Erlbaum Associates.