2005

Mitigating Weather Effects on the Future National Airspace System: the Integration of Human Factors, Decision Support and Display Technologies

Anthony D. Andre Ph.D.

Phil Smith Ph.D.

Amy Spencer Ph.D.

Jimmy Krozel Ph.D.

MITIGATING WEATHER EFFECTS ON THE FUTURE NATIONAL AIRSPACE SYSTEM: THE INTEGRATION OF HUMAN FACTORS, DECISION SUPPORT AND DISPLAY TECHNOLOGIES

Anthony D. Andre, Ph.D.
Interface Analysis Associates
1131 S. De Anza Blvd.
San Jose, CA 95129
andre@interface-analysis.com
408 966 9355

Phil Smith, Ph.D.
Cognitive Systems Engineering Lab, Ohio State University
Columbus, OH 43210
smith.131@osu.edu

Amy Spencer, Ph.D.
Cognitive Systems Engineering, Inc.
Ostrander, OH 43061
amyspencer@columbus.rr.com

Jimmy Krozel, Ph.D.
Metron Aviation, Inc.
Herndon, VA 20170
krozel@metronaviation.com
503-274-8316

Weather is a major limiting factor in the National Airspace System (NAS) today, accounting for roughly 65% of all traffic delays. Because we cannot control weather and safety must be maintained in the presence of weather- related hazards, our ability to mitigate the effects of weather through advances in weather prediction, human factors, decision support tools, automation and display technology are critical to supporting the projected growth in air travel demand. This paper presents the core ideas, human factors approach, and initial display concepts for supporting all-weather operations in the future NAS, developed as part of NASA's Virtual Airspace Modeling and Simulation (VAMS) program.

Introduction

Weather is a major limiting factor in the NAS today, accounting for roughly 65% of all traffic delays. Because we cannot control the weather and because safety must be maintained in the presence of weather related hazards, our ability to predict the weather and how it influences air traffic are critical elements in designing the future NAS. As part of NASA's Virtual Airspace Modeling and Simulation (VAMS) project we have been developing concepts for mitigating weather effects, and thus restoring or increasing the NAS capacity, for the years 2020 and beyond.

The capacity of the NAS is ultimately limited by its ability to accommodate safe and efficient travel under *all* weather conditions. The key to greater capacity in the NAS lies in our ability to accurately predict and adjust the future state of the NAS on a timescale consistent with critical NAS response times. From a Human Factors perspective we have developed a triad of core ideas to represent our concept for increasing the NAS capacity in the context of weather. The core ideas are: 1) flexible traffic management, 2) shared situation awareness, and 3) coupled weather and traffic prediction. The "Core Idea Triad" is based on the philosophy that the optimal plans, strategies and responses for mitigating weather effects cannot be fully achieved without common situation awareness among different NAS users, coordination of traffic plans, and sufficient information sharing and transfer.

We have developed a set of scenarios that depict both current day and future concept operations in the context of capacity-limiting weather events, across different levels of scope (e.g., local weather events, ground vs. upper air weather, propagating weather events) and involving different sets of NAS users (pilots, ATC, traffic managers, dispatchers, etc.). Each scenario details the weather phenomena in question, how the weather impacts current-day operations, future roles and responsibilities, decision support tools (DSTs) and other user interfaces derived from our core ideas and concepts. Further, for each scenario we have developed a preliminary set of functional illustrations, which serve to demonstrate the information that a given user, or set of users, might have access to in the future NAS.

A Human Factors Approach

A major element of this project was the identification of key human performance objectives, listed below. Our belief is that these issues underlie many of the previous failed attempts to introduce automation and decision aiding to the NAS at a large scale.

- Improve the Distribution of Data and Provide Tools to Assist with its Use.
- Make Technologies Useful in Spite of their Brittleness
- Constrain the Solutions Suggested by a DST Based on Human Factors Considerations.
- Support Collaboration and Coordination among distributed NAS operators.

A Human-Centered Design Process

As part of the concept design process, we first developed a high-level human-centered design approach. This approach is represented by the following main human factors themes.

- Implement new technology to enhance performance while employing human-centered design techniques to support human decision making, keep operators in the loop and in control.
- Utilize communication and display technologies to share relevant information and perspectives between pilots, dispatchers, controllers and traffic managers.
- Help formalize and automate useful procedures carried out today in a manual, effortful and ad-hoc manner.
- Maintain current roles and responsibilities as much as possible, but support proactive problem solving through advanced technology, human-centered DSTs and shared awareness interfaces.
- Develop realistic solutions that can implemented in the near-term or in phases over time.
- Design distributed work systems and procedures in order to avoid excessive cognitive complexity and workload for any one individual.

Our User Interface Approach

Our approach to user interface design, which we intend to apply to all operators within the NAS, is to impact the user's ability to access, understand, integrate, and act on the variety of information sources, and to do so in support of both individual and group work, in a timely fashion and with undue levels of workload and stress. New and emerging sensor, algorithm and display technologies will be considered in our effort. Finally, our interface design approach is supported by the following design principles.

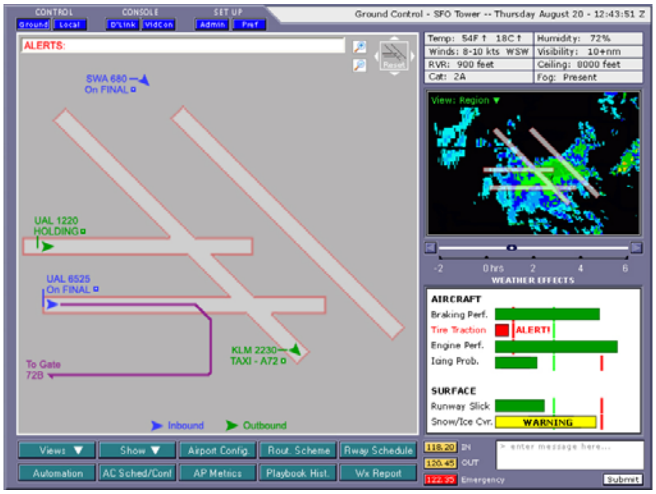
- Shared awareness – push relevant and context-sensitive, though not identical representations of, information to various NAS users towards facilitating collaborative decision making.
- User control/authority—support user, don't make decision for them.
- Transparency – allow the user access to the logic behind any calculation, algorithm or decision support solutions.
- Multi-modal – provide users with multiple information views or perspectives, taking advantage of different input and processing modalities.
- Collaborative – provide interfaces that make collaboration between NAS users efficient, easy and beneficial.
- Flexible – prevent automation and technology brittleness by allow the user to choose the parameters, to alter the logic, to add constraints not considered by the DST, to override automatically created values, and to adjust levels of uncertainty.
- Present Wx implications, not just data – provide the user with the implications or the effects of Wx, not just the data. In doing so, the interface is performing a common cognitive task for the human, that is, determining how Wx conditions will affect aircraft performance, airport surfaces, and other safety variables.
- Saliency – provide salient, at-a-glance indicators of overload, capacity loss, uncertainty, predicted effectiveness.
- Modeling and comparisons- provide the user with tools to model and compare various DST solutions, before selecting a specific initiative.
- User defined constraints – allow users to define and input constraints that may not be known to the computer system.
- Input of user priorities – allow timely and easy input and adjustment of user priorities.
- Visual modeling– provide layered visual representations of solutions, effects and Wx so that the user can easily see how the proposed

initiative will mitigate Wx and other constraints.

- History-provide the user with access to historical data (e.g., delayed or pop-up flights), success rates and system-derived estimates of the applicability of a DST solution to a given context or situation.

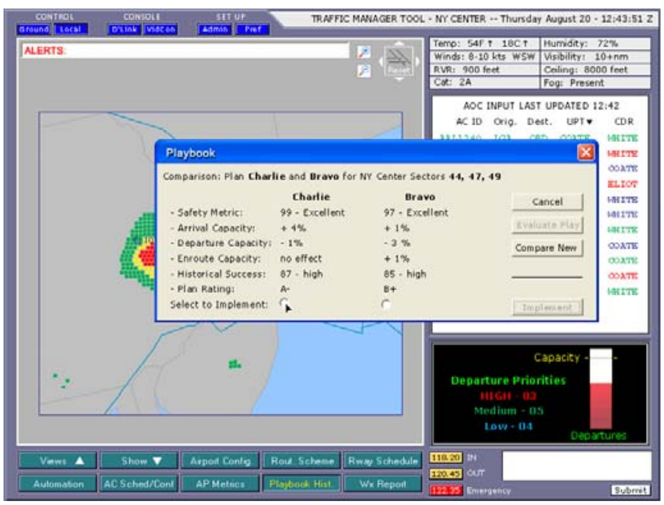
Concept Interfaces

On the following page we present some of the concept interfaces developed as part of this effort. They are already making a large impact on the aviation community and future plan for mitigating weather effects on our national airspace system.

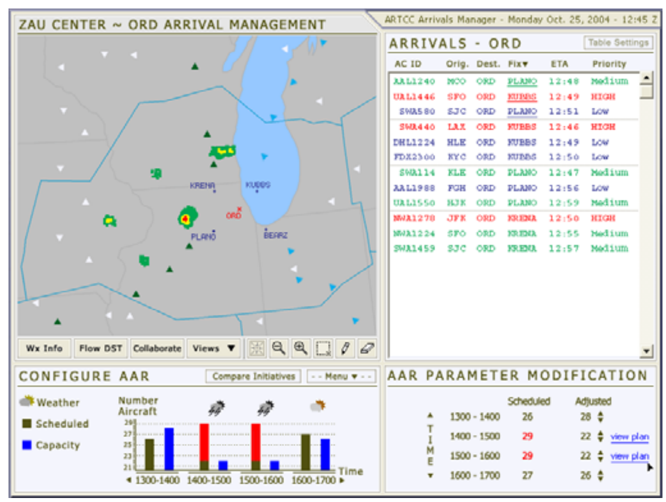


Future Ground Control SA Interface Showing Wx Effects

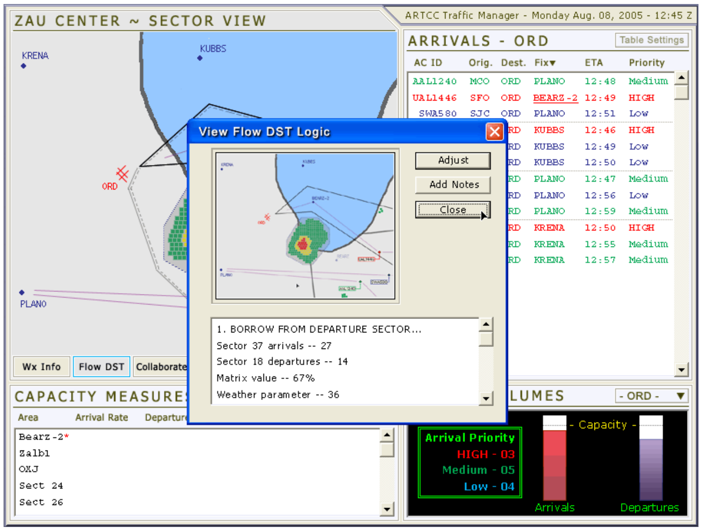
Future Ground Control SA Interface Showing Pilot View



Future Traffic Manager Interface Showing Plan Comparison



Future Center Controller Interface Showing Arrival Info



Future Center Controller Interface Showing DST Logic

Future Pilot Interface Showing Wx Optimized Route

References

Billings, C.E., "Toward a Human-Centered Aircraft Automation Philosophy", *International Journal of Aviation Psychology*, Vol. 1, No 4, pp. 261-270, 1991.

Hansman, J., and Davison, H.J., "The Effect of Shared Information on Pilot/Controller and Controller/Controller Interactions," *3rd USA/Europe Air Traffic Management R&D Seminar*, Napoli, June, 2000.

Krozel, J., Lee, C., and Mitchell, J.S.B., "Estimating Time of Arrival in Heavy Weather Conditions," *AIAA Guidance, Nav., and Control Conf.*, Portland, OR, Aug., 1999.

Penny, S., Krozel, J., Andre, T., *et al*, *Technologies Enabling All-Weather Maximum Capacity by 2020: Scenarios for the Capacity Increasing Concept*, Phase 2 Final Report, Technical Report, Metron Aviation, Inc., Herndon, VA, January 15, 2004.

LINE PILOT PERFORMANCE OF MEMORY ITEMS

Howard Au
Boeing Commercial Airplane Group*
P.O. Box 3707 MC 67-TC
Seattle, WA 98124, USA

* This research was paid for through out-of-pocket expenses by the author. It was conducted while the author was a student at Embry-Riddle Aeronautical University in partial fulfillment for the degree of M.S. Aeronautics.

An evaluation of Boeing 737 line pilot performance of memory items in 5 abnormal checklists was performed in a single-blind experiment using tabletop exercises at the crew base of a major U.S. airline. A study of 16 pilots shows that performance of memory items results in errors in identifying the failure, selecting the proper checklist to be completed, and checklist step errors.

Introduction

Some system failures that can occur on commercial airliners require flight crews to perform checklist steps from memory prior to referring to the checklist. These steps, called memory items or recall steps, are for time critical actions crucial to the safe continuation of the flight (e.g., preventing severe aircraft damage or crew incapacitation). Typically, line pilots do not study memory items except in preparation for a proficiency check (PC), usually every 6 or 12 months. They arrive for their evaluation prepared to be tested on the recall of the memory items. Their performance in these evaluations may not reflect their performance on the line, months after a PC.

This study examines whether line pilots are familiar enough with the memory items to perform all of them reliably, without prior knowledge that they will be evaluated. It was predicted that the performance of the memory items would show errors of commission, omission, and order due to the pilots' infrequent review of the memory items. This impromptu method of evaluation more closely resembles an unanticipated inflight emergency. This paper reviews some of the literature on performance under stress and then discusses the results pertaining to errors in identification of failures and errors in checklist selection. Although checklist step omission and order errors were observed, this paper will focus on the commission errors in the completion of checklist steps.

Human Performance Under Stress

An inflight emergency requiring timely action imposes a great deal of stress on the flight crew. Previous studies have shown that recall under high-stress conditions is more prone to errors than recall under low-stress conditions [8]. These errors, as they relate to checklist use, may include errors in identifying the abnormal condition, selecting the correct checklist, and errors of commission (adding steps or performing steps incorrectly), omission (missing steps), or order (completing steps in the wrong sequence).

Baddeley [1] presented a review of studies that included performance of deep-sea divers, combat aviators in actual combat, soldiers in simulated emergencies, and skydivers. These studies evaluated the performance of manual dexterity tasks, tracking tasks, and attention to peripheral cues. They showed that danger manifests itself in human performance through a narrowing of attention or through an increase in time to complete a manual dexterity task. The narrowing of attention can potentially lead to increased performance only if the task being performed is understood to be important. *However*, performance on tasks made to seem peripheral during an emergency can deteriorate [3]. Similarly, if the task is so complex as to require attention to numerous cues, the narrowing of attention will result in an inability to integrate relevant task information and an inability to conduct a proper assessment of the situation [6].

It is possible that training can mitigate some of these effects. *However*, even though pilots receive regular training in emergency procedures in simulators, that does not mean they are unaffected by the stress of an actual emergency. An emergency in a simulator is not perceived as life-threatening. If the pilot fails, the simulator can be reset for another attempt. Unless a pilot has had repeated experience in dealing with a truly dangerous emergency, performance in a real emergency could be similar to a novice. It has been shown that subjects are able to inhibit fear and prevent it from affecting their performance only if they are repeatedly exposed to a dangerous situation [1]. Due to the reliability of today's airliners, it is unlikely for the average airline pilot to have this kind of exposure in an airplane.

Stress Effects on Problem Analysis

It is possible that performance on infrequent tasks, such as identifying the root cause of multiple failures or shutting down an engine inflight, is affected more by stress than are common tasks. This is “an effect that has profound implications for the design of procedures to be used under the stressful conditions of emergency” [9].

This effect can sometimes be observed when people continue with a planned series of actions they are familiar with even when the actions appear unsuccessful or inappropriate. By acting before analyzing the situation, the operator may exacerbate the situation, which may induce more stress, and make it increasingly difficult to identify the original cause of the failure. This is related to an effect referred to as *confirmation bias*, where a person attends to cues that support a belief, and discounts cues that contradict the belief. Confirmation bias has been demonstrated in the use of automation and even in the diagnosis of everyday situations [4, 5, 7]. Other studies have shown that under stress, subjects are less effective and more disorganized at considering alternative solutions and incorporate less data in decision-making [6].

Stress Effects on Completion of Checklist Steps

Discussions with pilot participants in this study suggest that the requirement to perform certain actions from memory implies a sense of urgency in the performance of those actions. This introduces another potential source of error due to the loss of accuracy as speed is increased, an effect that is best described by the speed-accuracy operating characteristic (SAOC). The SAOC is a function that represents the inverse relationship between accuracy and speed. As the performance of a task requires more speed, accuracy is reduced until it approaches chance. If accuracy is excessively emphasized, then the time required to complete a task increases greatly with little improvement in accuracy.

Wickens & Hollands [9] summarize studies that demonstrate the effects of stress, induced by speed or by threat of bodily harm, on performance accuracy. For example, bomb-disposal experts performing under stress made more errors while working faster, and subjects who were threatened with the potential for electric shock gave up on problem-solving activities early.

Using an emergency descent as an example, an earlier study [2] showed that crews performing an emergency descent from memory took longer to descend than crews using the checklist. The difference in descent time resulted from omission errors by crews performing memory items. They occasionally omitted deploying the speedbrake, causing the airplane to descend slower. On the other hand, crews that performed the procedure by reference to the checklist did not make these errors, but took longer to complete the checklist. Regardless of the time required to read through the checklist, the crews performing the procedure by reference descended to a safe altitude in less time because of the use of the speedbrake.

The perceived requirement to perform checklist steps quickly from memory during high-stress situations is at odds with the need to perform those checklist steps accurately. There is a potential for loss of accuracy as the performance speed increases. Attempting 100% accuracy would require so much time to complete a checklist that other flying tasks would be disrupted. There is a tradeoff between getting the procedure done quickly, and getting it done while minimizing the possibility of error.

The following methodology seeks to identify examples of these errors in the flight operations domain. Even though inducing a level of stress similar to that of a real emergency was not possible in this study, it was hypothesized that errors of commission, omission, and order would still be observed.

Methodology

Participants

Sixteen 737 line pilots at a crew base of a major U.S. airline volunteered for the study. These pilots were already at the crew base either in preparation for a flight or returning from one. Participants were accepted without regard to experience level and participated in the study individually and not as a member of a two person crew. Pilots reported being trained in both the 737 Classic and 737 NG.

Procedure

In order to avoid any priming effects in the recall of their emergency procedures, subjects were not informed of the purpose of the research. They were instead briefed that:

- the research was on the suitability of the 737 alerting system,
- they would be asked to talk through five procedures, and
- the results from this study may be relevant to the design of a new alerting system in future airplanes.

A brief survey of experience was collected. This included data on total number of hours flown, their time in airplane type, flying time since last PC, and their crew position.

Subjects were seated in front of a poster of the flight deck. For consistency, a color poster of the 737 Classic flight deck was used. Five non-alerted abnormal procedures that contain memory items were used. They included aborted engine start, engine limit/surge/stall, rapid depressurization, runaway stabilizer trim, and dual engine failure.

The experimenter began each scenario by describing a normal flight situation, and then interjecting cues that suggest a particular failure. Subjects were asked to react to the cues as they would inflight, performing any procedures they felt were necessary. When responses to the scenarios seemed vague, the researcher probed the participants to encourage them to elaborate. The participants were provided with their airline Quick Reference Handbook (QRH), and were allowed to select the checklist they felt was most appropriate for the situation. Each session lasted approximately 30 minutes.

Results

Demographics

The participants in this study were 16 current line pilots at a major U.S. airline. Of those pilots, one was eliminated from the final analysis because he determined during the interview that an evaluation of the performance of memory items was the goal of the research.

	Total Time	Months Since PC	Time in Type	Weeks Since QRH Used
Mean	13,404	6	6,614	13
Standard Deviation	6,829	4	6,535	17
Minimum	4,500	0.5	400	1
Maximum	25,000	11	20,000	52

Table 1. *Demographics*

Data from the experience survey is presented in Table 1. Nine First Officers and six Captains participated. Two pilots incorrectly reported their total time and time in type, and their numbers were excluded. Seven pilots had prior military experience ranging from land and carrier-based fighters to large transports. Pilots who did not have military experience came from various corporate jets, commuter planes, other large commercial airlines, and corporate turboprops.

Checklist Selection Errors

When pilots were given an engine start condition with no oil pressure indications, four pilots initially chose the Engine Low Oil Pressure checklist. Upon reading that checklist, two of those pilots realized it was not appropriate for the situation, and correctly selected the Aborted Engine Start checklist. One pilot reported that there was no checklist needed, and that a maintenance call would be the only action required after completing the engine shutdown. The remaining 10 pilots correctly referenced the Aborted Engine Start checklist (Table 2).

Checklists selected	# of pilots
Aborted Engine Start	10
Engine Low Oil Pressure	2
Engine Low Oil Pressure > Aborted Engine Start	2
None	1

Table 2. *Aborted Engine Start Checklist Selection*

The Engine Limit/Surge/Stall scenario had the lowest identification rate (Table 3). Only two pilots referenced the correct checklist. One of those two selected the Engine Fire/Severe Damage/Separation checklist first. The remaining pilots referenced various checklists, including Engine Fire/Severe Damage/Separation, Engine Failure/Shutdown, and Engine Overheat.

Checklists selected	# of pilots
Engine Limit/Surge/Stall (Correct)	1
Engine Fire > Engine Limit / Surge / Stall (Experimenter prompted the correct checklist by saying the engine was “surging”)	1
Engine Failure	6
Engine Fire	4
Engine Overheat > Engine Fire	1
Engine Overheat	1
Engine Overheat > Engine Failure	1

Table 3. *Engine Limit / Surge / Stall Checklist Selection*

The remaining three scenarios had few checklist selection errors. One pilot selected the Auto Fail/Unscheduled Pressurization Change checklist during a rapid depressurization. Another pilot performed the Stabilizer Out Of Trim checklist in the runaway stabilizer scenario.

Checklist Step Errors

The majority of checklist step errors occurred during the completion of the dual engine failure memory items. Many of those were commission errors. These included:

- bringing the thrust levers back to idle before attempting to restart the engine,
- advancing the thrust levers as the engines failed in an attempt to get them to restart,
- starting the APU to try an assisted start,
- waiting three seconds to attempt a restart after shutting off the fuel,
- placing the ignition selector to both, and
- using engine anti-ice (Figure 1).

→ Ignition Selector................................Both

→ Thrust Levers..............................Advance

→ Engine Start Levers............................Idle

Engine Start Switches........................Flt

→ Turn around

→ Thrust Levers..................................Close

→ Engine Anti-ice.....................................On

Engine Start Levers......................Cutoff

EGT decreasing:

→ Wait three seconds:

Engine Start Levers...........................Idle

→ APU..Start

If EGT exceeds 950°C:

→ Allow engines to overheat

Repeat above steps

→ Attempt restart one at a time

Figure 1. *Dual Engine Failure Commission Errors. Bold items indicate the correct steps. Arrows indicate all additional steps performed by the 15 pilots.*

In the rapid depressurization scenario, two pilots included additional steps:

- verifying the engine bleeds were on, and
- closing the bleed air isolation valve (Figure 2).

→ Engine bleed switches..........................On

→ Isolation valve..................................Close

Oxygen masks & regulators....On/100%

Crew communications............Establish

Pressurization mode selector.........Man

Outflow valve................................Close

Figure 2. *Rapid Depressurization Commission Errors.*

Four pilots made commission errors in the completion of the runaway stabilizer trim checklist by attempting to activate the electric trim switches in the direction opposite the runaway. One of those four pilots stated that he would also attempt to engage a different autopilot in the hopes that it would not experience the same malfunction (Figure 3).

Control column.....................Hold firmly

Autopilot (if engaged)...........Disengage

→ Electric trim in opposite direction

→ Engage other autopilot

If runaway stabilizer continues:

Trim cutout switches...............Cutout

Trim wheel.....................Grasp & hold

Figure 3. *Runaway Stabilizer Commission Errors.*

Discussion

Checklist Selection Errors

When presented with cues to an abnormal situation, pilots sometimes omit a thorough analysis of the situation. This became evident through previous observations of pilots performing abnormal procedures in simulators and anecdotal evidence. The pilots in this study demonstrated a tendency to fixate on the most prominent cue and perform the checklist appropriate to that cue. However, a thorough analysis of the situation can reveal that the single most prominent cue does not always lead the pilot to the correct checklist.

There were 23 checklist selection errors. With the following three exceptions, the errors appear to be caused by the pilots' fixation on a single cue. Experimenter error in describing the rapid depressurization failure to one pilot gave the impression that the cabin altitude began to stabilize at approximately 12,000 feet, which led him to the Auto Fail/Unscheduled Pressurization Change checklist. Another error was due to a pilot's belief that no checklist was required for an aborted engine start. Finally, one pilot referred to the Dual Engine Failure checklist as the Engine Inflight Start checklist, but performed the correct memory items.

The remaining 20 checklist selection errors appear to be caused by pilots fixating on a single cue, and performing the checklist that appears most related to that cue. For example, in the aborted engine start, the cues given to the pilots were the continued illumination of the LOW OIL PRESSURE light and no oil pressure indication. Four pilots stated that, given those cues, they would complete the Low Oil Pressure checklist.

Two of those pilots realized the Low Oil Pressure checklist was inappropriate by considering the reasonableness of the checklist steps they were reading. The checklist directed the pilots to the Engine Failure/Shutdown checklist, which is meant for an inflight engine shutdown. A shutdown of an engine on the ground is simpler than a shutdown inflight and these pilots determined that irrelevant steps such as: starting the APU, maintaining fuel balance, and preparing for a single-engine landing, indicated they were in the wrong checklist. However, one pilot who entered the Engine Failure checklist from the Low Oil Pressure checklist did not consider the appropriateness of the checklist steps he was reading, and showed a tendency for perseveration. He went so far as to complete the Engine Failure checklist, reading aloud and bypassing irrelevant steps to complete the only step required to actually shutdown the engine while on the ground.

In the engine limit scenario, the 14 subjects who did not select the correct checklist instead performed the checklist that most closely reflected the cue they said was the most important. One pilot initially selected the Engine Fire/Severe Damage/Separation checklist, but turned to the Engine Limit/Surge/Stall checklist only after the experimenter said the engine was "surging". The term "surging" was not used as a cue in any other scenarios. Pilots who were primarily concerned by the abnormal "popping" or "banging" noises referenced the Engine Fire/Severe Damage/Separation checklist, stating that they believed the noises suggested severe engine damage. Pilots who considered excessive exhaust gas temperature (EGT) to be more important completed checklists related to overheat conditions. The pilot who referenced the Stabilizer Out Of Trim checklist in the runaway stabilizer scenario did so because he believed the STAB OUT OF TRIM light would be illuminated.

Checklist Step Errors

There appear to be consistent patterns in the observed checklist step errors. Many of the commission errors appear to result from the pilots' creativity in dealing with an abnormal situation. It was observed that many pilots perform steps in addition to what was required based on their understanding of how the airplane systems functioned, even though their understanding of the systems may be incorrect. Some pilots explained that the performance of some additional steps occurs because of knowledge of the intricacies of a complex system gained over years of experience or knowledge of common and simple failure modes, which are not addressed in the checklist. This may resolve the situation without the need for a checklist. In other cases, an incorrect or incomplete understanding of the system may lead pilots to perform additional steps that delay the completion of steps necessary to resolve the situation, or that may exacerbate the condition.

The pilots' creativity in dealing with certain situations was most evident in the dual engine failure scenario, which had the highest number of commission errors. A possible explanation was apparent in the pilots' response to this scenario: a desire to "do whatever it takes" to resolve a serious situation. Their perception was that this failure was so severe that they would exercise their authority as pilots, beyond what is written in the checklist, in an attempt to get an engine running, regardless of the consequences. Some pilots' willingness to allow the engines to exceed EGT and overheat, contrary to the guidance in the checklist, demonstrated this belief.

Most errors of commission were intended to troubleshoot the failures, such as: advance the thrust levers, verify the start levers are at idle, turn around to exit the heavy rain that caused the failure, and manually select both igniters. This last step demonstrates a misunderstanding of the ignition system. By correctly completing the recall item in the checklist, both igniters were automatically energized.

When the situation called for a shutdown of both engines, two pilots performed the additional step of delaying 3 seconds between restart attempts. They explained that this stemmed from a folk belief carried over from their military background that additional time was needed for excess fuel to clear the engine before attempting a restart.

This disposition towards creative troubleshooting was also seen in the Runaway Stabilizer Trim and Rapid Depressurization checklists. Errors of commission included moving the electric trim switches in the opposite direction and engaging the other autopilot. One pilot reported that he had experienced a runaway stabilizer in the past, and activating the electric trim switches stopped the runaway. This is an example of a pilot's knowledge of the failure modes of a complex system that could resolve the situation without using a checklist.

The rapid depressurization scenario showed that some commission errors, such as closing the isolation valve and ensuring the engine bleeds are on, would not exacerbate the situation, but would not be beneficial either. They would simply delay the completion of the necessary steps. Moreover, the manual closing of the isolation valve demonstrates a lack of understanding of the bleed air system. This step is not required because the valve is already closed during its normal operation.

On the other hand, some commission errors aggravated the situation. An example was seen in some pilots' willingness to allow the engines to overheat while restarting after a dual engine failure. The consequence of the overheating could be engine damage and a true engine failure, instead of the original problem of a temporary flameout due to an environmental condition such as heavy rain, resulting in no engine damage.

Conclusion

The results demonstrate that pilots have difficulty identifying the cause of the failure and selecting the correct procedure. After identifying the situation, knowledge of the appropriate memory items is such that pilots commit errors in recall even during unstressed conditions with a poster of the flight deck for context.

None of the five failure scenarios in this study had a distinct indicator light that would annunciate the condition. Pilots were forced to analyze the cues and determine the appropriate procedure. This is an uncommon and involved task, and not performing it may force pilots to complete only those tasks they are familiar with, such as following an illuminated LOW OIL PRESSURE light to the Low Oil Pressure checklist during an aborted engine start, or fixating on abnormal engine noises and performing the Engine Fire/Severe Damage/Separation checklist, instead of the more appropriate Engine Limit/Surge/Stall checklist.

The observed checklist step errors showed that pilots commit a number of errors. The majority of the commission errors were steps performed by pilots to resolve a failure based on their knowledge of the airplane systems. Some of these commission errors demonstrated a misunderstanding of how the systems in the 737 functioned. Other errors were a result of either knowledge gained during a real experience in the past, or a belief carried over from previous organizations and airplanes, which may no longer be applicable.

Implications

Even though the method used in this study did not induce stress, it allowed for an evaluation of the pilots' knowledge of the memory items without prior preparation. Pilots generally perform well during their PCs, and possibly better than inflight, because they expect an evaluation and can prepare for it. Pilot performance observed in this study may be closer to that in an inflight emergency, in which the pilots are unprepared to perform their memory items.

Clearly, an inflight emergency places a pilot under a great deal of stress. Based on the literature review, it can be inferred that errors similar to those observed here may occur inflight during an actual emergency, and may even occur more frequently due to increased stress. Conducting a similar study in a full-flight simulator may provide a level of stress similar to what is experienced in a real emergency. The results obtained from a simulator could be a more realistic representation of the results obtained inflight.

Acknowledgments

The author wishes to thank Barbara Holder and Randy Mumaw for their valuable guidance throughout the course of this project. The author also wishes to thank reviewers of drafts and the pilots who assisted by reviewing and commenting on research methodology, which greatly influenced the quality of data collected. The author also wishes to thank those who gave their time to allow us to gain some insight into the challenges they face in the cockpit.

References

1. Baddeley, A. D. (2000). Selective attention and performance in dangerous environments. *Human Performance in Extreme Environments*, 5(1), 86-91. (Reprinted from Baddeley, A. D. 1972. *The British Journal of Psychology*, 63(4), 537-546.)
2. Hamman, W. R. (n.d.). Quick Reference Checklist (QRC). *United Airlines validation study*. Colorado: Author.

3. Idzikowski, C., & Baddeley, A. D. (1983). Fear and dangerous environments. In D. H. Holding (Series Ed.) & G. R. J. Hockey (Vol. Ed.), *Stress and fatigue in human performance* (pp. 123-144). New York: John Wiley & Sons.

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

5. Reason, James. (1990). *Human Error*. New York: Cambridge University Press.

6. Salas, E., Driskell, J. E., & Hughes, S. (1996). Introduction: The study of stress and human performance. In E. A. Fleishman (Series Ed.) & J. E. Driskell & E. Salas (Vol. Eds.), *Stress and human performance* (pp. 1-45). New Jersey: Lawrence Erlbaum.

7. 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), 85-97.

8. Thompson, L. A., Williams, K. L., L'Esperance, P. R., & Cornelius, J. (2001). Context-dependent memory under stressful conditions: The case of skydiving. *Human Factors*, 43(4), 611-619.

9. Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. New Jersey: Prentice Hall