JOURNAL OF AIRCRAFT
Vol. 43, No. 5, September–October 2006

Human–Computer Interaction Analysis of Flight Management System Messages

Lance Sherry*
San Jose State University Research Foundation, Moffet Field, California 94035
Karl Fennell†
United Airlines, Denver, Colorado 80207
Micheal Feary‡
NASA Ames Research Center, Moffet Field, California 94035
and
Peter Polson§
University of Colorado, Boulder, Colorado 80309

DOI: 10.2514/1.20026

Researchers have identified low proficiency in pilot response to flight management system error messages and have documented pilot perceptions that the messages contribute to the overall difficulty in learning and using the flight management system. It is well known that sharp reductions in pilot proficiency occur when pilots are asked to perform tasks that are time-critical, occur very infrequently, and are not guided by salient visual cues on the user-interface. This paper describes the results of an analysis of the pilot human–computer interaction required to respond to 67 flight management system error messages from a representative modern flight management system. Thirty-six percent of the messages require prompt pilot response, occur very infrequently, and are not guided by visual cues. These results explain, in part, issues with pilot proficiency, and demonstrate the need for deliberate design of the messages to account for the properties of human–computer interaction. Guidelines for improved training and design of the error messages are discussed.

I. Introduction

ONE of the responsibilities for a pilot of a modern airliner is to respond to error and alert messages from the automation [1]. Researchers have raised concerns about pilot proficiency in response to these messages [2–4]. The Federal Aviation Administration (FAA) human factors team [5] has explicitly identified this area as a topic that requires further research.

The multifunction control and display unit (MCDU), illustrated in Fig. 1, serves as a passive user-interface for the flight management system (FMS). The scratchpad, shown in Fig. 1, provides a single-line display area for annunciation of messages. FMS error messages include the following categories:

1) system failures (e.g., SINGLE FMC L OPERATION)
2) mismatch between two cockpit systems (e.g., INERTIAL/ORIGIN DISAGREE, NAV INVALID TUNE XXX)
3) incompatible pilot entry (e.g., CHECK ALT TGT)
4) missing pilot entry (e.g., ENTER INERTIAL POSITION)
5) alert based on prediction (e.g., INSUFFICIENT FUEL)
6) reminders to pilot (e.g., DRAG REQUIRED, END OF OFFSET)
7) FMS capacity limits exceeded (e.g., ROUTE FULL)

Thirty-three percent of the FMS error messages occur in response to an entry made by the pilot. These messages are feedback for the task currently being performed by the pilot. The remaining 67% of the messages appear following an external event (e.g., system failure) and are not related to tasks currently being performed by the pilot. These messages may interrupt the pilot's current task and effectively create new tasks for the pilot. For example, the NAV INVALID TUNE XXX message requires the pilot stop what he or she is doing and determine why the navigation aid required for the selected approach procedure has not been tuned automatically.

Observations of revenue service and airline training have identified low proficiency in pilots' responses to FMS error messages [1,6,7]. The appearance of a message usually results in the questions "What does this mean?" and "What do we do about it?" In the absence of instructions embedded in the text of the message pilots are obliged to "explore" the user-interface, using cues provided in the message, the button labels, and page displays, to seek an appropriate response [8]. This behavior, known as "persistent interaction," draws a pilot's attention from other tasks and consumes cognitive resources [9]. The result is longer training periods, less efficient cockpit operations, and reduced margins of safety.

It is a basic principle of human–computer interaction (HCI) that human proficiency degrades rapidly when faced with unexpected events in time-critical environments [10]. Most cockpit error messages, by their nature, occur very infrequently at unexpected times and require prompt pilot response. Singer and Dekker [11] demonstrated that the design of the messages plays a significant role in pilot proficiency. They found statistically significant improvements in pilot proficiency when, in addition to alerting the pilot, the messages guide the pilot's next actions.

This paper describes the results of an analysis of the human–computer interaction required to respond to scratchpad error messages of a modern FMS. The major results are as follows:

1) Fifty-four percent of the messages are estimated to occur less than once in 100 flights. For 60% of the messages prompt pilot action may be appropriate to correct the condition.

Received 12 September 2005; revision received 6 January 2006; accepted for publication 9 January 2006.

*Senior Research Associate, Human Factors Research & Technology, NASA Ames Research Center, MS: 262-4, Moffet Field, CA 94035

†Senior Instructor, Boeing 777 Fleet, United Training Center, 7401 Martin Luther King Boulevard, Denver, CO 80207

‡Research Psychologist, Human Automation Integration Research, NASA Ames Research Center, MS: 262-4, Moffet Field, CA 94035

§Professor, Department of Psychology, 2690 Heidelberg Drive, Boulder, CO 80309

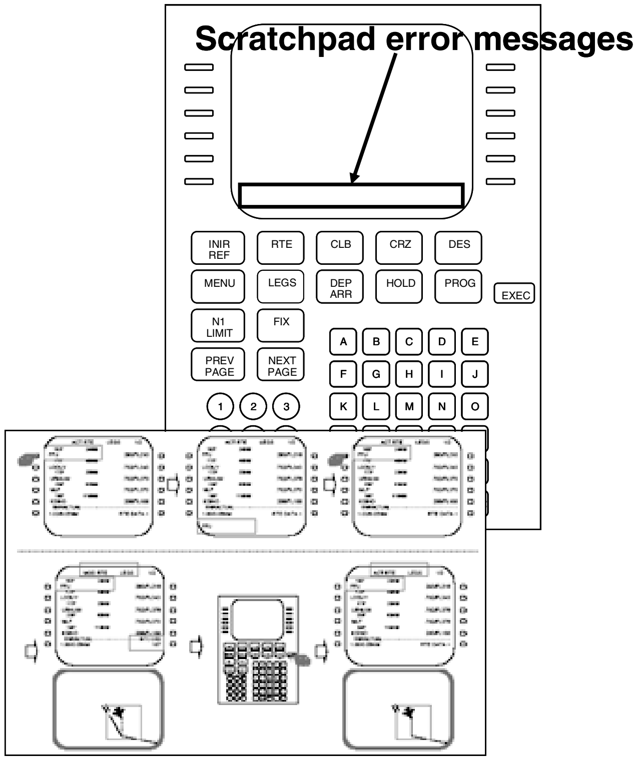

Fig. 1 MCDU is the user-interface for the FMS. Messages are displayed in the scratchpad. Inset shows example sequence of pilot actions required in response to a message.

2) Thirty-eight percent of the messages occur less than once in 100 flights and prompt pilot action may be appropriate to correct the condition.

3) Fifty-seven percent of the messages do not provide the pilot with the underlying causes or guide the next action.

4) Thirty-six percent of the messages are estimated to exhibit the preceding three properties: 1) occur less than once in 100 flights, 2) prompt pilot response may be appropriate, and 3) the message does not explicitly provide visual cues to guide the pilot response.

The combination of the low frequency of occurrence, the need for prompt pilot response, and the absence of visual cues explain, in part, the observed difficulties pilots experience in responding to the FMS error messages. These results also identify the need for deliberate design of the messages to ensure that they provide unambiguous visual cues (e.g., labels and prompts) to guide pilot response. This is an opportunity to reduce training costs, improve operational efficiency, and contribute to improved safety margins.

Section provides an overview of HCI for pilot response to the messages. Section III describes the method for analyzing the HCI response to messages. The results of the analysis are presented in Section IV. The paper concludes with design guidelines to optimize the HCI to respond to the messages in Section V. This paper is a summary of NASA TM-2005-213459 [12].

II. Overview of Human–Computer Interaction in the Cockpit

The pilot uses cockpit automation, such as the FMS, to aid in the execution of airline mission tasks, including instructions from air traffic control (ATC), airline standard operating procedures, checklist items, and responses to cautions/warnings and error messages from the automation. Some airline mission tasks can be performed directly by an automation function (e.g., Hold or Direct To). Other mission tasks must be reformulated by the pilot into a sequence of subtasks that use the functions of the automation. Other tasks cannot be supported by the functions of the automation and must be performed manually.

Pilot execution of the airline mission tasks using the automation involves a two-way communication between pilot and automation [1]. The operator communicates intentions to the automation using input control devices on the user-interface. The automation acknowledges pilot commands and provides feedback of its commanded behavior and the changes in the environment over time through the user-interface. This interaction requires a sequence of cognitive, physical, and perceptive actions in an execution and evaluation cycle (e.g., Norman [13]) as follows:

1) The pilot is required to decide what function of the automation to use to achieve the objective of the airline mission task.

2) The pilot decides on the sequence of button pushes to access the correct MCDU page.

3) The pilot then decides on the button pushes to enter the required data. Decision-making is required to determine the correct format, range, and location of entry of the data.

4) The pilot then confirms and executes the entry by seeing the appropriate changes in the MCDU and the cockpit displays (e.g., ND).

5) The pilot then monitors the automation and airplane to ensure that the airline mission objectives are achieved.

This five-step process is an idealized information flow model of the HCI required to execute an airline mission task [14,15]. Although there is no evidence that an expert pilot actually performs these five cognitive steps, the process is useful for identifying the type of perception, cognition, and motor skills for each step, the source of the cue (visual, inference, or memorized) that triggers each step, and the type of interaction and feedback that is involved.

For the purpose of an engineering analysis, the complex cognitive processes required to perform the action sequence for each of the five steps, is as follows. Each step is cued by either 1) a visual cue on the user-interface (e.g., label or prompt), 2) inferencing from user-interface conventions and similar tasks, or 3) a memorized cue. A visual cue is a label or prompt on the user-interface that cues the pilot to perform the next action. For example, the label on a button that matches the airline mission task cues the pilot on the availability of a function. On the other extreme, a memorized cue requires the pilot to recall from long-term memory the action for the HCI step. For example, the pilot recalls that the automation includes a "hidden" function that can perform the current airline mission task. Between the two extremes of visual and memory cueing, pilots reason out the required actions by using knowledge about user-conventions and knowledge for similar tasks. This is known as "performing by exploration" [8].

A. Time-to-Proficiency

The time-to-proficiency for a pilot learning to perform a task is directly proportional to the presence of salient visual cues that guide pilot actions, and to a lesser degree the consistency of user-interface conventions across tasks, and the similarity between tasks within a group. Prompts and labels on the user-interface provide constant reminders of the appropriate actions and rapidly strengthen the cognitive associations that result in appropriate response. The absence of labels and prompts requires the user to accrete knowledge of user-interface conventions and similarity of tasks by experience. When this information is not explicitly provided in the training material or covered during training, it can take a long time to identify and build the knowledge structure [13].

In a recent study of the performance of 20 C-130 pilots qualified within the previous six months on an advanced FMS [16], the likelihood of a pilot not being able to access the correct FMS page was 74% if the there were two pilot actions that were not prompted by visual cues. The likelihood of a pilot not being able to access the correct FMS page was 13% when there was a single pilot action not supported by visual cues. In contrast, when the pilot actions were visually cued, the likelihood of a pilot not accessing the correct page was 6%.

B. Proficiency in Revenue Service

The proficiency in executing a task in revenue service is directly proportional to the frequency with which a task is performed [10]. Repetition results in cognitive processing that occurs without mental effort, is stress tolerant, and is not corrupted by interruptions [17].

For tasks that are not performed frequently, proficiency is determined by the presence of visual cues, the consistency of the user-interface conventions between tasks, and the similarity with other tasks [18]. The absence of these cues and reliance on memorized action sequences results in reduced pilot proficiency [19].

Singer and Dekker [11] demonstrated this principle in a study of pilot proficiency in commercial airline operations in responding to caution/warning messages (such as Loss of All Engines, or Loss of Hydraulic Pressure). Messages classified as “Sort & Guide,” in which the message provided salient visual cues (e.g., prompts) to guide pilot response, exhibited significantly better performance in response time and error rates than messages that did not provide guidance.

III. Method

The purpose of this study was to analyze the HCI required to respond to FMS error messages. The messages were analyzed to determine the frequency of occurrence, the operational impact of the message, and the presence of salient visual cues to guide pilot response to the message.

The messages were classified and analyzed to criteria by a team of four experts that included one flight instructor with 16 years of experience at a major U.S. airline, one senior cognitive scientist with over 40 years of experience, one multiengine jet-rated pilot and flight instructor who is also a human factors researcher with 10 years of experience, and one avionics design engineer with 18 years of experience designing these systems. All classifications of the tasks were made by complete concurrence of all parties through a process of dialogue. The small number of discrepancies between individual assessments was due to domain aeronautical knowledge. It should be noted that the estimate of frequency of occurrence is the most subjective result reported. This estimate is biased by the experiences of the subject-matter-experts and the type of flying (line or simulator) they are exposed to. Despite this constraint, the accuracy of this estimate could be considered relatively high due to the broad ranges used (see next section). Based on this experience, we believe another analysis conducted by this team (or by another team with similar levels of experience) would yield the repeatable results within an acceptable margin.

The Boeing 777 was selected as a representative modern FMS. The 67 messages analyzed in this study were all of the messages listed in Section 8 of the Boeing 777 Flight Management System Pilot’s Guide [20]. The results of this study can be generalized to all FMSs designed for commercial airliners.

A. Estimate of Frequency of Occurrence

An estimate of frequency of occurrence was identified for each message. The estimate was based on revenue-service flight experience and jump-seat observations. The categories are as follows:

1) Very infrequent: occurs once in every 101 + flights
2) Infrequent: occurs once in every 21–100 flights
3) Occasional: occurs once in every 5–20 flights
4) All: occurs once in every 1–4 flights

B. Operational Impact of the Event That Results in a Message

An estimate of the operational impact of the event that results in each message was made. The estimate accounted for the worst-case condition that prompted the message. The assessment was made based on revenue-service flight experience and knowledge of some of the design criteria. The actual design specifications of the conditions in the software were not available for this study.

The categories for worst-case operational impact are as follows:

1) Prevents planned flightplan: prohibits execution of flightplan filed (e.g., diversion, delayed gate pushback). These conditions could result in significant excess costs to the airline. Prompt pilot action may be appropriate.

2) Prevents expected procedure: prohibits execution of anticipated procedure (e.g., RNAV approach). An alternate procedure can be used but requires additional pilot workload or additional procedure (e.g., go-around). These conditions could result in additional costs to the airlines (e.g. hold, go-around). Prompt pilot action may be appropriate

3) No impact to flightplan or procedure: no impact to flight. Pilot may address message as time permits.

It should be noted that in several cases, the occurrence of a message was deemed to be a symptom of a more significant operational condition than the software conditions under which the message appears. For example, the FUEL DISAGREE message appears when “Totalizer (TOTL) fuel quantity and FMC computed (CALC) fuel quantity disagree by 9000 pounds for more than 5 minutes” [20]. Airline operational experience indicates that this is not merely a disagreement between two parameters that generally are within some range, but could be an indicator of a possible fuel leak (resulting in execution of a 30 min fuel leak procedure), a failure of the fuel quantity indicating system, or a failure of the engine fuel flow sensor. For the purpose of this study the worst-case condition was used.

C. Presence of Salient Visual Cue to Guide Pilot Actions

This analysis identifies the type of cue (see, infer, or remember) that leads the pilot to perform the actions in response to the message. A message is categorized as *see* when the phraseology of the message leads directly to the action sequence. A message is categorized as *infer/remember* when the phraseology of the message requires the pilot to 1) use knowledge of interface conventions, 2) use knowledge of similar tasks, or 3) retrieve knowledge from long-term memory to interpret and/or determine the action sequence. For example, the message TAKEOFF SPEEDS DELETED is informative but does not lead to the appropriate pilot action to return to the takeoff page and confirm the V-speed calculations. The pilot has to build proficiency in training to remember that the response to this message is to access the takeoff page and reconfirm the takeoff speeds.

IV. Results

The results of the analysis are summarized in the next section. A complete list of the messages and the complete analysis is included in NASA TM-2205-213459 [12].

A. Frequency of Occurrence

The estimate of frequency of occurrence for the messages was determined. Thirty-six of the messages (54%) were considered to occur very infrequently (once in more than 101 flights). An example of a very infrequent message is FUEL DISAGREE—PROG 2/2, indicating a mismatch in fuel sensors. Eleven messages (16%) were considered to occur infrequently (once in more than 20 flights, but less than 100 flights). An example of an infrequent message is NAV INVALID—TUNE XXXX, indicating that the Navaid in the published RNAV or VOR approach is currently not being tuned. The low repetitions of these tasks are likely to result in low pilot proficiency when pilot action is not guided by visual cues.

Nineteen messages (28%) were considered to occur occasionally (once in every 5–20 flights). These included messages in response to pilot typo entries INVALID ENTRY, NOT IN DATABASE, and INVALID DELETE. One message, TAKEOFF SPEEDS DELETED, occurs on every flight because of the incompatibility between airline operational flows and the sequence of entries required by the FMS.

B. Operational Impact of Event

An assessment was made of the worst-case event that results in each message. Three messages (5%) were classified as appearing due to an event that prevents planned flight. These events require the pilot to interrupt his or her current tasks and respond to the message. The FUEL DISAGREE—PROG 2/2 message would prompt the pilot to execute the fuel leak checklist: a 30 min procedure that could lead to a diversion and manual landing. The other two messages occur during preflight at the gate and prevent on-time pushback from the gate: CHECK AIRLINE POLICY and NAV DATA OUT OF DATE.

Forty messages (60%) were classified as appearing in response to an event considered to prevent an expected procedure. These events require the pilot to interrupt their current tasks and respond to the message. Failure to respond immediately does not jeopardize the planned flight or result in a failure to meet near-term ATC instructions. For example, RUNWAY N/A FOR ARR requires the pilot to investigate the reason why the runway and/or approach selected by the pilot is not compatible with the pilot-selected standard arrival procedure). TAKEOFF SPEEDS DELETED requires the pilot to reconfirm the V-speeds following a modification to the FMS flightplan or other FMS parameters before takeoff.

Twenty-four messages (36%) were classified as appearing in response to an event with no flightplan or procedure impact. These events do not require immediate attention but should be addressed by the pilot on a time-available basis. Example messages include END OF ROUTE and UNABLE CRZ ALT XXX.

C. Frequency of Occurrence and Operational Impact of Event

The relationship between frequency of occurrence and operational impact of event is summarized in Table 1. Thirty-eight percent of the messages occur very infrequently and prevent execution of the filed flightplan or an expected procedure. A total of 54% of the messages occur very infrequently or infrequently and prevent execution of the filed flightplan or an expected procedure This has strong implications for the way the messages must be worded and trained. Because of the long interval between observations of a given message, it must be assumed by the designer that the pilot will not be able to recall the appropriate response to the message. As a consequence it would be prudent for the designer to ensure that the messages cue the pilot to the next action. This is discussed further in the Conclusions section.

D. Presence of Salient Visual Cue to Guide Pilot Actions

Next, the analysis evaluated whether the content of the message provides the pilot with a visual cue (e.g., prompt) to perform the next action, or whether the pilot must infer or remember the next action. Twenty-nine messages (43%) were supported by salient visual cues. Thirty-eight messages (57%) were not supported by salient visual cues and had to rely on inferencing or remembering.

E. Frequency, Operational Impact, and Visual Cueing

The combination of the three properties described in the preceding section determines the degree to which the design of the messages facilitate proficiency in learning and using the system.

Thirty-six percent of the messages (i.e., 24 out of 67 messages) were estimated to occur very infrequently, to represent an event for which prompt pilot response is appropriate, and do not guide pilot response with salient visual cues.

Based on the research of pilot human–computer interaction, it is very likely that these messages will increase the pilot's workload and require persistent interaction. This result has strong implications on how these messages should be trained. This result also suggests the necessity to explicitly design pilot human–computer interaction to ensure that these designs are avoided in the future.

V. Conclusions

The results of this analysis demonstrate that the design of the FMS messages do not explicitly take into consideration the properties of human cognition in the presence of infrequent and time-constrained tasks. This directly contributes to the perceptions held by pilots that the FMS is hard to learn and difficult to use.

Frequency of occurrence is the primary determinant of a pilot's ability to reach proficiency in performing the action sequences required in response to a message. High-frequency occurrence of a message ensures high cognitive activation levels that allow pilots to reliably perform these tasks even in the presence of interruptions. For tasks that are performed infrequently, visual cues, such as labels and prompts, are required to guide the pilot to the appropriate action sequence. It should be noted that visual cues also provide feedback to catch slips in executing frequent tasks and provide the scaffolding to streamline learning the required HCI.

The results of the analysis of the messages indicate that, by their nature, a large percentage of FMS messages will occur infrequently and require rapid, robust pilot response. The application of the HCI design principles requires that the messages provide visual cues to lead the pilot to perform the correct response. The visual cues in the messages need to exhibit two characteristics:

1) Identify the situation, the resulting task, and the next HCI action. One format could be <situation>-<task>-<next action>.

2) Use terminology already in use on the MCDU page titles, labels, and data.

For example, the existing message INSUFFICIENT FUEL describes a situation in which the FMS predictions indicate that the usable fuel on board will be less than the pilot entered reserves at the destination. This situation occurs when the cruise winds used by the predictions do not reflect the cruise winds planned by the airline dispatch, or when the FMS flightplan represents a significantly longer route (due to an error in waypoint or procedure entry). A redesigned message would read: FUEL AT DEST < RESERVES—CHECK WINDS/LEGS/RTE. The message explicitly refers to the terminology used on the PERF INIT page in the description of the situation that leads to the message. The pilot task prompted is to verify the entry of the winds, the legs, and the route.

Examples of redesigned messages are included in the table in Appendix A. in [12].

One of the criticisms of this proposal is that the new messages require additional space on the display. In many cases the proposed messages include more characters than can be displayed on the singleline scratchpad that is currently used. Researchers and designers have proposed a "pop-up dialog box" for these longer error messages (e.g., Abbott [21]). The result is a design tradeoff between increased usability and possible additional display clutter. When the cost of pilot training and some measures of cockpit operational efficiency are included in the tradeoff, the additional design, implementation, and tests costs are relatively small.

To ensure quality in the design of FMS messages and all cockpit automation HCI, Sherry and Feary [14] have proposed processes and documentation compatible with standard industry engineering practices (e.g., DO-178B software development process). The kernel of this process is the task design document (TDD) that provides a design specification for HCI required to perform all airline mission tasks. In this way the pilot–avionics interaction is not an emergent property of design process, but a deliberate design decision that takes into consideration the effect of the user-interface on training and use.

To provide the basis for a design and testing of the HCI, the TDD includes 1) a definition of the tasks, 2) the mapping between tasks and

Table 1 Relationship between frequency of occurrence and operational impact

Operational impact	Frequency			
	Very infrequent	Infrequent	Occasional	All the time
Prevents flightplan	2	1	——	——
Prevents procedure	24	9	5	1
Not flightplan or procedure	9	1	13	——

intended functions of the avionics equipment, 3) the sequence of steps used by the pilot to interact with the avionics to perform each task, and 4) the source (see, infer, or remember) for each step.

Airline mission tasks that are complex to execute and mission tasks that require excessive inferencing or memorization of action sequences become apparent in this specification. Redesign of the function or the user-interface may be required. Alternatively, tasks that require inferencing and memorization of action sequences may require a "waiver" from the program manager before the TDD review can be signed off. The TDD can also be used for credit in the FAA human factors certification process.

The results of the study identify shortcomings in the design of the messages that are not likely to be corrected in the existing fleet. The workaround is to build and maintain the cognitive skills of pilots to respond to the infrequent messages that require prompt pilot response.

There are two elements to this workaround. First, the pilot must be provided with training material that includes a complete set of pilot actions to be performed in response to each message. These instructions should explicitly highlight the visual cues for actions based on visual cues, the inferencing rules for actions based on user-interface conventions or similar tasks, or the memorization items for actions that must be memorized. The five-step process described in Section may be a useful way to organize the action sequences. Second, the pilots must be provided opportunities to maintain proficiency for the infrequent tasks. Mechanisms include web-based quizzes, written multiple choice quizzes, and captain/first-officer oral quizzes. The latter could be an airline procedure completed during cruise on a revenue-service flight.

Acknowlegdments

This work was supported by NASA Aviation Operations Systems Program contract GS09T01BHM0386 order ID 9T1N001H (technical point of contact: Michael Feary). Technical contributions were made by Immanuel Barshi (NASA-ARC), Mike Matessa (NASA-ARC), Kevin Jordan (SJSU Research Foundation), Paul Schutte (NASA-LaRC), Dennis O'Donoghue (Boeing, now Eclipse Aviation), Randy Mumaw (Boeing), Dan Boorman (Boeing), Steve Quarry (Honeywell), Anne Coldwell (Medtronic), William Easto (Smiths Aerospace), Michael DeJonge (Smiths Aerospace), Melanie LeGoullon (George Mason University), Michelle Harper (University of Central Florida), Kevin O'Neill, Guy Woolman (Southwest Airlines) and the three anonymous reviewers.

References

[1] Billings, C. E., *Aviation Automation: The Search for a Human Centered Approach*, Lawrence Erlbaum Associates, Mahwah, NJ, 1996.

[2] Hicks, M., and deBrito, G., "Civil Aircraft Warning Systems: Who's Calling the Shots?," *International Conference on Human Computer Interaction in Aeronautics (HCI-Aero '98)*, edited by G. Boy, C. Graeber, and J. M. Robert, Ecole Polytechnique de Montréal, Montreal, May 1998, pp. 205–212.

[3] Billings, C. E., and Dekker, S. W. A., "Advanced and Novel Concepts for the Future System," *Aviation Automation: The Search for a Human Centered Approach*, edited by C. E. Billings, Lawrence Erlbaum Associates, Mahwah, NJ, 1996, pp. 221–231.

[4] Woods, D. D., "Cognitive Demands and Activities in Dynamic Fault Management: Abduction and Disturbance Management," *Human Factors of Alarm Design*, edited by N. Stanton, Taylor & Francis, London, 1994. pp. 88–107.

[5] Federal Aviation Administration, "Human Factors Study Team Report on the Interface Between Flightcrews and Modern Flight-deck Systems," Federal Aviation Administration, Washington, DC, 1996.

[6] "BASI: Advanced Technology Safety Survey Report," *Flight Safety Digest Special Issue*, Flight Safety Foundation, June–Aug. 1999, pp. 137–216.

[7] Martensson, L., and Singer, G., "Warning Systems in Commercial Aircraft: An Analysis of Existing Systems," Royal Institute of Technology, TRITA-IEO-1998:01, Stockholm, 1998.

[8] Polson, P., and Smith, N., "The Cockpit Cognitive Walkthrough," *Proceedings of the Tenth Symposium on Aviation Psychology*, Ohio State University, Columbus, OH, 1995, pp. 427–432.

[9] Norman, D., *Learning and Memory*, W. H. Freeman and Company, New York, 1982.

[10] Wickens, C. D., *Engineering Psychology and Human Performance*, HarperCollins, New York, 1992.

[11] Singer, G., and Dekker, S. W. A., "Pilot Performance During Multiple Failures: An Empirical Study of Different Warning Systems," *Transportation Human Factors*, Vol. 2, No. 1, 2000, pp. 63–76.

[12] Sherry, L., Fennell, K., Feary, M., and Polson, P., "A Human–Computer Interaction Analysis of Flight Management System Messages," NASA TM-2005-213459, 2005.

[13] Norman, D., *The Design of Everyday Things*, MIT Press, Boston, 1988.

[14] Sherry, L., and Feary, M., "Task Design and Verification Testing for Certification of Avionics Equipment," *Proceedings of AIAA/Digital Avionics Systems Conference*, Vol. 2, IEEE, Piscataway, NJ, Oct. 2004, pp. 10.A.3-100–10.A.3-110.

[15] Polson, P., Irving, S., and Irving, J. E., "Application of Formal Models of Human–Computer Interaction to Training and Use of the Control and Display Unit," Final Report, System Technology Division, ARD 200, Department of Transportation, Federal Aviation Administration, Washington, DC, 1995.

[16] Fennell, K., Sherry, L., Roberts, R., Feary, M., "Difficult Access: The Impact of Recall Steps on Flight Management System Errors," *International Journal of Aviation Psychology*, Vol. 16, No. 2, 2006, pp. 175–196.

[17] Shiffrin, R. M., and Schneider, W., "Automatic and Controlled Processing Revisited," *Psychological Review*, Vol. 91, No. 2, 1984, pp. 269–276.

[18] Lewis, C., Polson, P., Wharton, P., and Rieman, J., "Testing a Walkthrough Methodology for Theory-Based Design Evaluation of Walk-Up-and-Use User-Interfaces," *Proceedings CHI* , ACM, New York, April 1990, pp. 235–242.

[19] Javaux, D., "Explaining Sarter & Woods Classical Results: The Cognitive Complexity of Pilot-Autopilot Interaction on the Boeing 737-EFIS," *Proceedings of the Second Workshop on Human Error, Safety, and Software Design*, edited by N. Leveson and C. Johnson, Univ. of Washington, Seattle, WA, 1997, pp. 62–77.

[20] Air Transport Systems Division, *Boeing 777 Flight Management System Pilots Guide*, Honeywell, Phoenix, AZ, 1995.

[21] Abbott, T., "A Comparison of Two Control Display Unit Concepts on Flight Management System Training," NASA TM-4744, 1997.