Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)

METHODOLOGY FOR ESTIMATION OF BENEFITS OF HUMAN-FACTORS ENGINEERING IN NEXTGEN/SESAR DEVELOPMENT

Lance Sherry (Ph.D.)
Center for Air Transportation Systems Research
George Mason University
Fairfax, VA., USA
lsherry@gmu.edu

Michael Feary (Ph.D.)
NASA Ames Research Center
Moffet Field, CA, USA.
Michael.feary@nasa.gov

Jerome Lard (Ph.D.)
Thales Research & Technology
Cedex, France
Jerome.lard@thalesgroup.com

Capt. Karl Fenell
United Airlines
Denver, CO
KarlFennell@ual.com

***Abstract*— The concepts-of-operation proposed for Next Generation Air Transportation System (NextGen) and Single European Sky ATM Research (SESAR) implicitly require a significant improvement over existing standards for flightdeck human-computer interaction. Whereas in today's airspace operations there is no routing penalty for delayed response to a required maneuver, flights in high density NextGen/SESAR airspace that are unable to respond to off-nominal situations in a timely manner, will lose their slot and be shifted to a downgraded level of airspace resulting in flight delays and/or increased route distance.**

The flightdeck automation required for NextGen/SESAR concepts-of-operations is estimated to be equivalent in magnitude to the introduction of the Flight Management System (FMS) to the flightdeck starting in the 1980's. Researchers documenting the introduction of the FMS in revenue-service operations identified widespread difficulties in certification, training, and use (or lack of use). It is likely that the patterns of misuse, disuse, and abuse of flightdeck automation will be repeated, unless explicit design interventions are made in the NextGen/SESAR development. To justify the additional expense for the explicit design of the human-computer interaction during the NextGen/SESAR development process the benefits of cost savings in airline revenue-service must be established.

This paper describes a methodology for estimating the revenue-service cost savings that can be derived from HCI Engineering in the development of NextGen/SESAR flightdeck automation. An example of the cost savings benefits accrued by a hypothetical large U.S. domestic carrier due to the redesign of FMS error messages ($45M per year) is provided along with a discussion of the implications and limitations of the cost savings model.

Keywords-cost/benefit analysis; human-computer-interaction, flightdeck automation, NextGen/SESAR concepts-of-operations

I. INTRODUCTION

Concepts-of-operations proposed for the Next Generation Air Transportation System (NextGen) and Single European Sky ATM Research (SESAR) are designed to increase the effective airspace and airport capacity [1, 2]. This will be achieved largely by more efficient use of the airspace through trajectory-based operations and reduced separation in high-density airspace.

Explicit in these concepts-of-operations is increased navigation performance *as well as* increased flightcrew proficiency in performing the required maneuvers in a timely and robust manner. Whereas in today's air traffic control paradigm, flights that do not meet the required performance do not lose their access to preferential airspace, in NextGen/SESAR, flights that cannot respond to off-nominal situations in a timely manner will be shifted to downgraded airspace and experience flight delays and/or extended distance routing.

Meeting the criteria for robust human-computer interaction to ensure seamless operations is not guaranteed. It is estimated that the additional functions to be the added to the flightdeck to support NextGen/SESAR will be roughly equivalent in magnitude to the upgrade to flightdeck automation that occurred with the introduction of Flight Management Systems (FMS) during the 1980's. Researchers extensively documented issues with the certification, training, and use (or lack of use) of the FMS in revenue service operations [3, 4, 5, 6, 7]. To avoid the pattern of misuse, disuse, and abuse [8], new practices in the design, such as HCI Engineering, of the flightdeck automation functions must be adopted.

There are three main roadblocks to the widespread deployment of emerging HCI Engineering tools for NextGen/SESAR. First, the benefits are accrued by one stakeholder (i.e. the airlines), while the burden of the expenses is shouldered by other stakeholders (i.e. aircraft manufacturers and avionics vendors). Second, quantifiable benefits of HCI Engineering are confounded with the performance of the overall enterprise including airline operations, airline training, and even air traffic congestion and weather. Third, there is widespread skepticism concerning the benefits of HCI

Engineering in the development community. This skepticism is best described as the “Cold Water and Empty Guns” syndrome documented by Rouse [9] in a seminal paper on the value-added of human factors. Rouse determined that the managers of large, complex development programs perceived the contributions of human factors specialists to be: (i) “throwing cold water” on design proposals due to concerns for the human factors and, (ii) being unable to mitigate these concerns with alternate design proposals (i.e. “empty guns”).

The landscape for HCI Engineering has changed radically over the last decade. First, increasing airline cost pressures and low-cost business models have reached a tipping-point that has incentivized manufacturers to invest in HCI Engineering during product development [10, 11]. Second, advances in “affordable” operator performance models [12, 13, 14] have enabled the development of a new breed of tools for automated analysis of proposed automation[15, 16, 17, 18, 19, 20]. These tools will overcome the “cold water, empty guns” reservations. However, the missing piece of the puzzle is a means to quantify the direct benefits of HCI engineering for the stakeholders in a way that decouples the effects of efficiency in human-computer interaction from the overall enterprise operations.

The purpose of this paper is to describe a methodology for the estimation of the revenue service cost savings to airlines resulting from the investment of HCI Engineering during the development phase of NextGen/SESAR automation. The paper is organized as follows: Section 2 provides an overview of HCI and the specific phenomenon of “persistent interaction” that reduces operational efficiency and safety margins. Section 2 also describes a Flightcrew Performance Model (FPM) used to predict the HCI persistent interaction phenomenon. Section 3 describes a first-order cost model for assessing the benefits of the investment of HCI persistent interaction. This cost savings model is based on the FPM described in the previous section. Section 4 includes an example case-study that describes the benefits that could be accrued by a hypothetical large domestic carrier resulting from the HCI engineering of FMS error messages ($45M per year). Conclusions, implications, limitations and future work are described in Section 5

II. HUMAN-COMPUTER INTERACTION AND PERSISTENT INTERACTION

This section provides an overview of the human-computer interaction required to complete mission tasks using flightdeck automation and describes the phenomena of *persistent interaction*.

A. Human-Computer Interaction

Interaction between the human operator and automation on the flightdeck of a commercial airliner occurs when the flightcrew use the automation to complete mission tasks. Mission tasks are the explicit set of tasks required to conduct the mission and include actions such as: executing ATC instructions, performing checklist items, and responding to caution/warning messages.

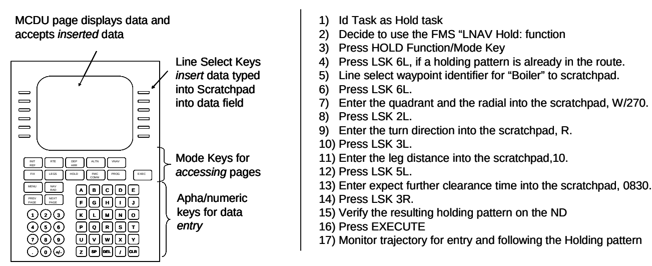
These tasks, initiated by voice communication, aural alerting, visual cues, or prospective memory, are performed using flightdeck automation such as the Flight Control Computer and its Mode Control Panel (MCP) or the Flight Management System (FMS) and its Multi-Function Control and Display Unit (MCDU). Table 1 provides a sample of mission tasks performed using the MCP and MCDU.

Mission Tasks triggered by ATC:	Mission Tasks triggered by FMS Error Messages:
“Climb maintain FL 2-2-0” “Proceed direct to waypoint XXX” “Intercept course 0-1-4 degrees from present position” “For weather, you are cleared to offset your current route 20 nautical miles right. Report when clear of the weather” “Fly heading 180, join the Gordensville 060 degree radial, track the radial inbound to Gordensville” ”Hold west of Boiler on the 270^{O} radial. Right turns. 10 mile legs. Expect further clearance at 0830”	Diagnose a mismatch in fuel sensors triggered by message FUEL DISAGREE – PROG 2/2 Diagnose error in automated procedure tuning triggered by message NAV INVALID – TUNE XXXX – Diagnose flightcrew data entry problems triggered by messages: INVALID ENTRY, NOT IN DATABASE, and INVALID DELETE Re-enter data required before takeoff triggered by message: TAKEOFF SPEEDS DELETED –

Sample of mission tasks. Some mission tasks are generated by ATC, other mission tasks are generated by checklists and error/caution/warning messages.
Table 1.

When the task is triggered, the flightcrew must determine which function of the automation to use. For example, to execute the ATC instruction to perform a Hold at a waypoint, the flightcrew should use the FMS “LNAV Hold” function which is explicitly designed to automate the task of flying holding patterns with an accuracy that cannot be achieved flying manually or by using the MCP (especially in the presence of a crosswind). In some cases, mental math is required to convert parameters from the ATC instruction to data that can be entered into the automation (e.g. reciprocal of course).

After deciding which function to use, the flightcrew must *access* the function. In the case of the MCDU this may involve several button-pushes to locate the correct page. Then the flightcrew must *enter* the appropriate data or make the appropriate selections. In most cases, the entry/selection must be “cross-checked” with the other pilot and then “executed” by



Example interaction with MCDU required to execute task for Hold west of Boiler VOR on 270 degree radial.
Figure 1.

an additional operator action (e.g. EXECute key on the MCDU). For many tasks, especially those associated with following the flightplan, the airplane trajectory must be monitored closely. A summary of the operator actions for the task "Hold west of Boiler on the 270^{O} radial. Right turns. 10 mile legs. Expect further clearance at 0830" is shown in Figure 1. For more information on canonical models of human-computer interaction see Sherry et.al [21, 22].

B. Properties of Mission Tasks

Not all mission tasks are created equal. Two functional properties that distinguish mission tasks are *frequency* and *impact on mission*.

Each mission task is performed with different *frequency*. Tasks that are performed routinely on every flight (e.g. pre-flight the FMS) are classified as Frequent. Occasional tasks are performed once every 5 flights (e.g. offset lateral path). Infrequent tasks are performed approximately once every 20 flights or less. Tasks that occur less than once every 100 flights are considered Very Infrequent. These tasks are typically associated with abnormal or emergency procedures (e.g. engine-out).

The *impact on the mission* reflects the value of the task to the conduct of the mission. On the one extreme, mission tasks classified with impact of Flight Critical, such as an error message indicating a mismatch between fuel estimates and actual fuel levels, must be addressed immediately for continued safe flight. Mission tasks classified with impact of Procedure Critical, are required to be performed for safe flight and/or yield efficiency benefits (e.g. runway/approach selected in Flightplan is not compatible with the arrival in the flightplan). Mission tasks classified as Neither Flight/Procedure Critical have no safety or efficiency implications and can be performed at flightcrew discretion.

Table 2 summarizes the frequency and impact on the mission for 67 error messages from a B777 FMS [23]. The typical error message profile is weighted towards the infrequent occurrence. Forty-six out of 67 messages are infrequent or very infrequent. The majority of the messages are flight critical or procedure critical. The estimated values of importance for each classification of Mission Impact are also shown in the table. These values are the estimated average flight delay experienced by a flight that is delayed in executing a mission task.

C. Persistent Interaction

The interaction between the flightcrew and automation can be classified into categories: (1) flawless (or smooth) interaction in which the flightcrew execute the most direct sequence of decision-making and button-pushing actions to complete the task, and (2) flawed interaction in which the flightcrew, unsure what action to take next *use trial-and-error exploration* of the user-interface to identify the correct sequence of button pushes [24]. The use of trial-and-error exploration of the user-interface, ubiquitous across all domains in HCI, is observed as *persistent interaction* by the flightcrew. That is, the operator is devotes their full attention to the interaction to keep track of the labels and pages visited in the process of discovering the correct action sequence. This tunneling effect results in degraded flightdeck efficiency and reduced safety margins.

	Frequency			
Mission Impact (estimate of additional flight time if fail to complete the mission task)	Very Infrequent (> 100 flights)	In-frequent (= 20 flights)	Occasional (=5 flights)	All the Time
Flight Critical (+30 mins delay)	2	1	-	-
Procedure Critical (10 mins delay)	24	9	5	1
Not Flight or Procedure Critical (< 1 mins delay)	9	1	13	-

Distribution of mission tasks generated from 67 error messages for the B777 FMS. See[23]
Table 2.

An example of documented persistent interaction in the flightdeck and the role it played in reduced safety margins is described in the analysis of the American Airlines Flight 965 accident at Cali, Columbia [25]. There were several phenomena and circumstances that contributed to the accident, but as Ladkin documented, the flightcrew spent an inordinate amount of precious time heads-down trying to manipulate the FMS to perform an infrequent task which is not well supported by the interface (underlined text below).

> "Confusion about the FMS presentation, as is true for use of any computer, is often resolved after *persistent interaction* with it. Thus it is likely that the Captain believed that the confusion he was encountering was related to *his* use of the FMS, and that continued interaction with it would ultimately clarify the situation."

> "He could not be expected to recognize, because it rarely occurs in regular operations, that the fix he was attempting to locate (Tulua VOR) was by this time behind him, and the FMS generated display did not provide sufficient information to the flightcrew that the fix was behind the airplane."

> "… because of the lack of time, the need to perform multiple tasks in the limited time, … his attention thus narrowed to further repeated unsuccessful attempts to locate ULQ [Tulua VOR] through the FMS."

D. Flightcrew Performance Model for Persistent Interaction

The Flight 965 anecdote points to two main factors that determine flightcrew performance in completing mission tasks: frequency of occurrence, and the availability of salient visual cues on the user-interface. Analysis of human-computer interaction of mission tasks in commercial airline flightdecks identified that mission tasks that are performed frequently tend to be properly completed in a timely manner. The frequent repetition of these tasks ensures that flightcrews will remember the required mental math or sequence of button-pushes to complete the tasks.

As the frequency of use drops-off, the performance of flightcrews is determined by a second factor – the degree of visual cues that guide the flightcrew to perform the next action. A flightcrew's ability to complete mission tasks for low frequency and rare tasks that are not supported by prompts or visual cues relies entirely on the recall of memorized action sequences. Recall of long action sequences from memory is not reliable [26, 27].

The phenomenon of persistent interaction experienced by flightcrews can be reduced significantly by the explicit design of mission tasks and the labels and prompts that guide operator actions to complete the task [21]. This is the intervention in the NextGen/SESAR development process that is required to eliminate HCI persistent interaction.

III. COST SAVINGS MODEL

The benefits of employing HCI Engineering for the analysis of persistent interaction is derived directly from the reductions in flight delays and in excess distance flown in revenue service when flightcrews are unable to flawlessly use the automation to accomplish the full range of mission tasks. Estimates of the cost saving benefits accrued by flawless interaction with the automation are a function of the frequency of occurrence of a task, the probability of failure to compete the task (within a time threshold), and the cost of the additional flight time or distance flown.

An estimate of the cost savings benefits can be derived in the three step process as follows:

Step 1: Compute an Estimate of the Annual Excess Costs to Airline Revenue Service Operations:

$$\text{Excess Costs per Flight} = \sum \text{FOC(i)} * \text{PFtC(i)} * \Delta\text{FT(i)} * \text{ADOC}$$

where:

FOC(i) = Likelihood of Task i being performed on a given flight (or Frequency of Occurrence)

PFtC(i) = Probability of Failure to Complete Task i in time t_i

ΔFT(i) = Additional Flight Time Incurred by Failure to Complete the Task i (based on Mission Importance)

ADOC = Airline Direct Operating Costs

Daily Excess Costs to Airline = Excess Costs per Flights * Average Number of Airline Flights per Day

Annual Excess Costs to Airline = Daily Excess Costs to Airline * 365 days per year

Step 2: Adjust the PFtC(i) profile to account for redesign of the interface (through HCI Engineering of Persistent Interaction)

The goal of HCI Engineering for analysis of persistent interaction is to identify and fix tasks that exhibit the potential to have high values (e.g. > 0.3) for the Probability of Failure to Complete (PFtC) the task. Experience in conducting HCI Engineering for persistent interaction indicates that it is not always possible to eliminate the causes of high values of PFtC for all tasks. There are several reasons that a perfect PFtC cannot be achieved: differences in the "semantic state-space" held by each operator, inadequate physical space on the user-interface to list all the appropriate labels to match the task semantics, constraints on the space and format of labels and prompts. Preliminary research is underway to better quantify the variances in semantic state-space held by pilots and to develop techniques to group similar functions used by a family of tasks.

For these reasons, for the purpose of cost savings estimation, practical experience suggests that starting with a distribution of PFtC, incremental improvements can be made to move 10% to 25% of the tasks that are in excess of a PFtC threshold of 0.3 to less than 0.3. In general, designers will start with improving the performance for the tasks that exhibit the highest importance to the mission. An alternative approach for the purpose of cost savings analysis (.e.g. Section IV below), is to spread the improvement of PFtC equally amongst the classes of importance assigning the improvements to individual tasks on a random basis.

Use the same equations described in Step 1.

Step 3: Compute the Benefits Yielded from HCI Engineering of Persistent Interaction.

Take the difference between the results of Step 2 and Step 1.

Estimated Annual Cost Savings = Annual Excess Costs to Airline (Baseline) - Annual Excess Costs to Airline (HCI Engineering Applied)

The example in the next section demonstrates the application of the cost model.

IV. EXAMPLE BENEFITS ANALYSIS FOR HCI ENGINEERING OF FMS ERROR MESSAGES FOR HYPOTHETICAL LARGE DOMESTIC AIRLINE

The implementation of NextGen/SESAR functions to support RNP and RNAV procedures and 4-D trajectory contract negotiations will require the addition of automation error messages reflecting the status of navigation sensors, navigation precision, data-communications messages etc. It is assumed that the inability of the flightcrew to respond to the error messages effectively and in an efficient manner will result in reduced access to preferred airspace resulting in flight delays and/or excess distance flown.

For the purpose of estimating revenue service cost savings benefits, the distribution of the frequencies and the mission importance for the NexTGen/SESAR messages is assumed to exhibit the same profiles as the previous generation of FMS messages (see Table 2). Further, in the absence of HCI Engineering in the NextGen/SESAR development processes, the salience of visual cues on the interface for each message will follow similar patterns as the previous generation of messages. Specifically, 57% of the new messages will not provide flightcrews the underlying causes of the message or guide the subsequent flightcrew actions to respond to message, and that overall, 36% of the new messages will occur very infrequently, exhibit high mission importance, and not be supported by salient visual cues on the interface. For further information on the distributions of the FMS error messages see [23].

In addition to the assumptions above, the Total Flight Delays, Total Excess Costs per Day, and Total Excess Costs per Annum for a hypothetical large domestic carrier are calculated with the following inputs:

- Average Flights per Day – 3745 (535 aircraft that average 7 flights per day)
- Average hours utilization per aircraft – 13 hours with average flight distance – 630 nautical miles
- Average Direct Operating Cost - $35 per minute. Biased by ratio of gate-taxi airborne operations (Bureau of Transportation Statistics [30]).

Table 3 shows the estimate of the flight delays and their associated excess costs accrued by the airline due to persistent interaction with the error messages. The baseline for the study, shown in the left column, is for error messages that exhibit distributions of frequency, importance, and visual cueing similar to the existing FMS message set. The subsequent columns show the excess costs accrued by the airline when 10% and 25% of the error messages with PFtC greater than 0.3 are improved to a PFtC of less than 0.3.

For the airline route properties described above, the hypothetical airline could accrue between $122K and $306K in daily savings from the 10% and 25% reductions in Probability of Failure-to-Complete tasks triggered by these error messages. Annual savings range from $44M to $112M for the 10% and 25% improvement conditions.

V. CONCLUSIONS

This paper describes a methodology to estimate savings in Flight Delays and in Airline Direct Operating Costs (ADOC) achieved by efficient use of the functions of flightdeck automation. In particular the savings are generated: (1) by reduced flight-time (including fuel-burn and crew costs) that occur when flightcrews are able to respond to ATC instructions, checklist items, and caution/warning/error messages in a timely manner, and (2) when flightcrews are able to respond to ATC requests to expedite operations or control accurately to a crossing restrictions using advanced functions in flightdeck automation. Both of these circumstances result in reduced distance flown and greater flexibility in on-time performance (i.e. slowing down is better than having to make up time). The underlying assumption of this approach is that NextGen and SESAR airspace will be operated on the basis of equipage and the ability to perform required tasks. Flights that are unable to perform the required maneuvers will be shunted to alternate airspace and will be required to absorb additional costs.

The methodology described computes the expected value for cost savings based on the: (i) frequency of occurrence of the

	Baseline (Traditional Design Methods)	**HCI Engineering yields 10% less Tasks with PFtC > 0.3**	**HCI Engineering yields 25% less Tasks with PFtC > 0.3**
Daily Total Flight Delays	583 hrs	525 hrs (-58 hrs)	437 hrs (-146 hrs)
Daily Total Excess Cost	$1.2M	$1.1M ($-0.122M)	$0.92M ($-0.306M)
Annual Total Excess Costs	$447M	$402M ($-44M)	$335M ($-112M)

Summary of Total Excess Cost per Day and Total Excess Cost per Year resulting from additional flight times accrued by inefficient response to new messages to support NextGen/SESAR procedures. The new messages exhibit historic distributions for frequency and importance to mission, but through HCI Engineering, exhibit reduced Probability of Failure-to- Complete
Table 3

task, (ii) average cost of failing to complete the task (in minutes flight delay), and (iii) the Probability of Failure-to-Complete the task. Case studies illustrate the application of the cost model to airline revenue-service operations and shows that annual savings for a large domestic carrier can be estimated conservatively between $40M and $110M

A. Limitations & Future Work

The benefits model described in this paper is "system engineering" or "first-order model." The accuracy of the results of this model are dependent on the validity of the inputs: frequency of occurrence, mission importance, flight delays and excess distance flown. In the case-study in this paper, data for these inputs were estimated using a Delphi procedure with Subject-Matter-Experts (see [21]). Collecting actual data from the flightdeck is more difficult.

The predictions for Probability of Failure-to-Complete a task were based on an empirical operator performance model (see [22]). There are two approaches to improve the model. First, the model was based on data from two experiments that included twenty pilots performing twenty tasks [28] and ten pilots performing twenty tasks [11]. Additional subjects performing a wider range of tasks would improve the empirical model.

Second, the operator performance model computes the Probability of Failure-to-Complete a task (PFtC) based on the number of operator actions supported by label following. The assessment of the visual and semantic salience of these cues was assessed manually using a Delphi process and Subject-Matter-Experts. Research is underway to automate the assessment of salience using an Aviation Knowledge Corpus and Latent Semantic Analysis engine [14]. This automation would compare the semantics of the task description with the semantics of the labels and prompts on the user-interface and determine the salience (e.g. score from 0 to 1). See [12], [14].

Other research is underway to use the empirical operator performance models to calibrate simulation-based models. Simulation-based models, derived from models of cognition such as ACTR [29] and linked to the Aviation Corpus have the enable evaluation of operator performance outside the scope of the empirical data.

It is important to realize that the estimates in savings in excess costs are conservative (i.e. low) for four reasons. First they are based on a conservative (low) estimate of ADOC that is biased to include a 30/70 split between ground and airborne operations to account for delays taken on the ground and at the gate. Second, the savings in excess costs do not include any multiplier effects of airlines operations when delays impact connecting passengers and flightcrews. Third, this paper has not discussed the savings accrued in flightcrew training and qualification. Explicit design of tasks for reliable performance for infrequent use has the side benefit of also reducing the number of trials required to mastery. This has the potential to reduce both fixed and variable costs of airline training. The cost savings accrued through improved training and procedure design is the subject of further research. Fourth, this estimate does not take into account any of the benefits accrued through improved margins of safety.

The savings in excess costs estimates derived from this methodology should only be applied to the first year following installation. Subsequent years would yield a learning effect that would reduce the cost savings as the flightcrew's knowledge improved.

Another application of the cost savings model is the estimation of environmental benefits derived through reduced fuel-burn achieved seamless operations. This estimation would be useful in an environmental cap-and-trade paradigm in which emissions taxes are assessed and carbon-footprints are measured.

Finally, the phenomenon of persistent interaction can be reduced significantly by the explicit design of mission tasks and the associated operator actions and the labels and prompts that guide operator actions. The emerging class of HCI tools used in a modified DO-178 software development process (see [21] make this feasible. The cost of deployment of these tools has not been discussed in this paper and is the subject of future work.

ACKNOWLEDGMENT

This project was funded by a grant from NASA – Aeronautics – Intelligent Integrated Flightdeck program, by internal George Mason University Foundation Funds, and by THALES IR&D funds. Thank you for technical assistance and suggestions from: Dan Boorman, Randall Mumaw (Boeing), Karl Fennel (United Airlines), John Otiker (SWA), Peter Polson, Marilyn Blackmon (University of Colorado), Andrea Knight, Ricardo Prada (Google), Karlin Toner (FAA), Robert Mauro (University of Oregon), Immanuel Barshi (NASA-ARC), Mike Valencia, Midori Tanino (FAA), Mike Matessa (Alion Inc.), Michael Gordon (CCPace), Ken Goldwasser (SEER Tech), Tarik Ward (NASA-JSC), Chip Hathaway, Terry Thompson (Metron Aviation), Debra Shreckenghost (TRACLabs), John Shortle, Rajesh Ganesan, John Ferguson, Guillermo Calderon-Mesa, Jianfeng Wang, Loni Nath (CATSR/GMU). Thank you for support of the research to Steve Young (NASA), Amy Pritchett (NASA).

REFERENCES

[1] JPDO (2008) NextGen Integrated Work Plan Version 1.0. http://www.jpdo.gov/iwp.asp

[2] SESAR Consortium (2008) SESAR Master Plan D5. MGT-0803-001-01-00. https://www.atmmasterplan.eu/http://prisme-oas.atmmasterplan.eu/atmmasterplan/faces/index.jspx

[3] Wiener, E., D. Nagel (1989) *Human Factors in Aviation*. Academic Press : NY, NY.

[4] Federal Aviation Administration. (1996) Human factors study team report on the interface between flightcrews and modern flightdeck systems. Washington, D.C. 1996.

[5] Billings, C.E. *Aviation automation: The search for a human centered approach*. 1996. Lawrence Erlbaum Associates: Mahwah, N.J, Inc

[6] BASI: Advanced Technology Safety Survey Report. *Flight Safety Digest Special Issue.* Flight Safety Foundation, June-Aug 1999, pp. 137-216. Available at www.basi.gov.au/sprog/advtek.htm (9/00)

[7] Feary, M. M. Alkin, P. Polson, D. McCrobie, L. Sherry, E. Palmer (1999) Aiding Vertical Guidance Understanding. Air & Space Europe Volume 1, Issue 1, January-February 1999, Pages 38-41.

[8] Parasuraman, R., V. Riley (1997) Humans and Automation: Use, Misuse, Disuse, Abuse. Human Factors, Vol. 39, 1997

[9] Rouse W.B. (1986) Cold Water and Empty Guns: A Report from the Front. Presentation to the Department of Defense Human Factors Engineering Technical Advisory Group. Goleta, California

[10] Boorman, D.J., R. J Mumaw (2007) A New Autoflight/FMS Interface: Guiding Design Principles Air Canada Pilots Association, Safety Conference, 2007 acpa.ca

[11] Prada, L. Ricardo; Mumaw, R J.; Boehm-Davis, D. A.; Boorman, D.J. (2007) Testing Boeing's Flight Deck of the Future: A Comparison Between Current and Prototype Autoflight Panels Human Factors and Ergonomics Society Annual Meeting Proceedings, Aerospace Systems , pp. 55-58(4)

[12] Teo L., B.E. John B. E.(2008) Towards a Tool for Predicting Goal-directed Exploratory Behavior In *Proc. of the Human Factors and Ergonomics Society 52nd Annual Meeting*, 950-954

[13] Matessa, M. & P. Polson (2005) List Model of Procedure Learning**.** *NASA/TM 2005-213465*

[14] Blackmon, M. H., Kitajima, M., & Polson, P. G. (2005). Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. CHI 2005: Conference on Human Factors in Computing Systems (pp. 31-40). New York: ACM Press

[15] John, B. E. & Salvucci, D. D. (2005) Multi-Purpose Prototypes for Assessing User Interfaces in Pervasive Computing Systems. *IEEE Pervasive Computing 4* (4), 27-34.

[16] Feary, M. (2007) Automatic Detection of Interaction Vulnerabilities in an Executable Specification. Engineering Psychology and Cognitive Ergonomics Volume4562/2007 Springer Berlin / Heidelberg. pg 487-496.

[17] Ockerman, J. A. Pritchett (2004) Improving performance on procedural tasks through presentation of locational procedure context: an empirical evaluation. Behaviour & Information Technology, Volume 23 , Issue 1 (January-February 2004. pgs 11-20

[18] Savidis A., Stephanidis C., Unified user interface design: Designing universally accessible interactions (2004) Interacting with Computers, 16 (2), pp. 243-270

[19] J.L. Pérez-Medina, S. Dupuy-Chessa and A. Front (200&) A Survey of Model Driven Engineering Tools for User Interface Design . In Task Models and Diagrams for User Interface Design, Volume 4849/2007Springer Berlin / Heidelberg

[20] J. Lard, F. Landragin, O. Grisvard, D. Faure: Un cadre de conception pour réunir les modèles d'interaction et l'ingénierie des interfaces CoRR abs/0807.2628: (2008)

[21] Sherry, L. & M. Feary. Task design and verification testing for certification of avionics equipment. In *Proceedings of AIAA/Digital Avionics Systems Conference, 2004*. Salt Lake City, Utah. October, 2004

[22] Sherry, L., M. Medina, M. Feary, J. Otiker (2008) Automated Tool for Task Analysis of NextGen Automation. *Proceedings of the Eighth Integrated Communications, Navigation and Surveillance (ICNS) Conference (2008), Bethesda, MD.* May 5-7, 2008

[23] Sherry, L., K. Fennell, M. Feary, P. Polson (2006) Human–Computer Interaction Analysis of Flight Management System Messages. Journal of Aircraft. Vol. 43, No. 5, September–October 2006

[24] Polson, P., & Smith, N. The cockpit cognitive walkthrough. *Proceedings of the Tenth Symposium on Aviation Psychology.* Ohio State University: Columbus OH. 1995**.** pp. 427-432

[25] Ladkin, P. AA965 Cali Accident Report, Near Buga, Colombia, Dec 20, 1995. Peter Ladkin Universität Bielefeld, Germany. Available 6/12/04 at http://sunnyday.mit.edu/safety-club/calirep.html

[26] Wickens, C.D. *Engineering psychology and human performance*. Harper Collins: New York. 1992

[27] Norman, D. *The design of everyday things.* MIT Press: Boston. MA. 1988

[28] Fennell, K., L. Sherry, R. Roberts. Accessing FMS functionality: The impact of design on learning. NAS/CR-2004-212837. 2004

[29] Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

[30] Bureau of Transportation Statistics (2008) http://www.bts.gov

AUTHOR BIOGRAPHY

Lance Sherry (Ph.D ’99, M.Sc. ’95, B.Sc. ’84) is Associate Research Professor. of System Engineering and Operations Research and is Executive Director of the Center for Air Transportation Systems Research (CASTR) at GMU. Dr. Sherry is a system engineer with over 25 years of practical experience in air transportation operations and the design/flight-test/certification of commercial avionics. Dr. Sherry has served as control engineer, system engineer, lead system engineer, avionics flight test engineer, and program manager, has also served as Principal Investigator on research projects for FAA, NASA, NSF, DOT, DOE, airports, airlines, aircraft manufacturers and avionics vendors and has published over 100 papers and articles. He holds several patents and has won several awards for his work.

.