APPENDIX E

HUMAN-PERFORMANCE MODELS

This appendix comprises two summary papers describing the work on developing human-performance models that was conducted under the System Wide Accident Prevention (SWAP) project of the Aviation Safety Program, and some summary comments from the meeting session at which they were presented.

The first paper (Foyle, Goodman, and Hooey, (2003)) is the introduction to a conference proceedings and provides an overview of the activities, the state of modeling human performance at the end of 2002, and a description of the future activities planned for 2003 and 2004. It appears in the proceedings of the NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, *NASA Conference Proceedings NASA/CP-2003-212267*, 2003.

The second paper is a summary paper describing the modeling efforts that was presented as a panel presentation on the topic at the 2005 Human Factors and Ergonomics Society meeting. It appears in the *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Santa Monica, Calif., 2005.

AN OVERVIEW OF THE NASA AVIATION SAFETY PROGRAM (AVSP) SYSTEMWIDE ACCIDENT PREVENTION (SWAP) HUMAN PERFORMANCE MODELING (HPM) ELEMENT

David C. Foyle
NASA Ames Research Center
Allen Goodman
San Jose State University
Becky L. Hooey
Monterey Technologies, Inc.

Abstract

An overview is provided of the Human Performance Modeling (HPM) element within the NASA Aviation Safety Program (AvSP). Two separate model development tracks for performance modeling of real-world aviation environments are described: the first focuses on the advancement of cognitive modeling tools for system design, while the second centers on a prescriptive engineering model of activity tracking for error detection and analysis. A progressive implementation strategy for both tracks is discussed in which increasingly more complex, safety-relevant applications are undertaken to extend the state-of-the-art, as well as to reveal potential human-system vulnerabilities in the aviation domain. Of particular interest is the ability to predict the precursors to error and to assess potential mitigation strategies associated with the operational use of future flight deck technologies.

HPM Element Goals

This report provides a summary review of recent research activities conducted in support of the Human Performance Modeling (HPM) element within the Systemwide Accident Prevention (SWAP) Level 2 project of the NASA Aviation Safety Program (AvSP). In March 2003, a one-day conference was held at NASA Ames Research Center to present the interim results of the HPM element. Specifically, the 2003 NASA HPM conference was focused on scenarios related to approach and landing with synthetic vision systems (SVS).

The overall 5-year goal of the HPM element is to develop and advance the state of cognitive modeling while addressing real-world safety problems. To this end, the HPM element continues to develop and demonstrate cognitive models of human performance that will aid aviation product designers in developing equipment and procedures that support pilots' tasks, are easier to use, and are less susceptible to error. The modeling focus is on computational frameworks that facilitate the use of fast-time simulation for the predictive analysis of pilot behaviors in real-world aviation environments.

Rationale

More than two-thirds of all aircraft accidents are attributed to pilot error. Identifying when equipment and procedures do not fully support the operational needs of pilots is critical to reducing error and improving flight safety (Leiden, Keller & French, 2001). This becomes especially relevant in the development of new flight deck technologies, which have traditionally followed a design process more focused on component functionality and technical performance than pilot usage and operability. To help counter this bias and to better understand the potential for human error associated with the deployment of new and complex systems, advanced tools are needed for predicting pilot performance in real-world operational environments.

As noted in the literature on aviation safety, serious piloting errors and the resultant accidents are rare events (for a review, see Leiden, Keller & French (2001)). The low-probability of occurrence makes the study of serious pilot errors difficult to investigate in the field and in the laboratory. These errors characteristically result from a complex interaction between unusual circumstances, subtle "latent" flaws in system design and procedures, and limitations and biases in human performance. This can lead to the fielding of equipment, which puts flight safety at risk, particularly when operated in a manner or under circumstances that may not have been envisioned or tested.

When combined with nominal and off-nominal scenario human-in-the-loop real-time testing, human-performance modeling in non-real–time (usually, fast-time) simulations provides a complementary technique to develop systems and procedures that are tailored to the pilots' tasks, capabilities, and limitations. Because of its use in fast-time simulations, human-performance modeling is a powerful technique to uncover "latent design flaws" -- in which a system contains a design flaw that may induce pilot error only under some low-probability confluence of precursors, conditions and events

Human performance modeling using fast-time simulation offers a powerful technique to examine human interactions with existing and proposed aviation systems across an unlimited range of possible operating conditions. It provides a flexible and economical way to manipulate aspects of the task-environment, the equipment and procedures, and the human for simulation analyses. In particular, fast-time simulation analyses can suggest the nature of likely pilot errors, as well as highlight precursor conditions to error such as high levels of memory demand, mounting time pressure and workload, attentional tunneling or distraction, and deteriorating situational awareness. Fast-time simulation is the only practical way to generate the very large sample sizes that are needed to reveal low-rate-of-occurrence events. Human-in-the-loop real-time simulations are too costly to use for this purpose. Additionally, this can be done early in the design cycle, without the need to fabricate expensive prototype hardware.

HPM Models

The AvSP HPM element is organized along two model development tracks (see figure E-1). The first model development track is Predictive Human Performance Models, in which multiple predictive models of human performance simultaneously address several well-specified problems in aviation safety. The second model development track is the Prescriptive Engineering Human Performance Model, which consists of a model of error detection (specifically a prescriptive engineering model of operator performance in context). The six models comprising the AvSP HPM element are described below.

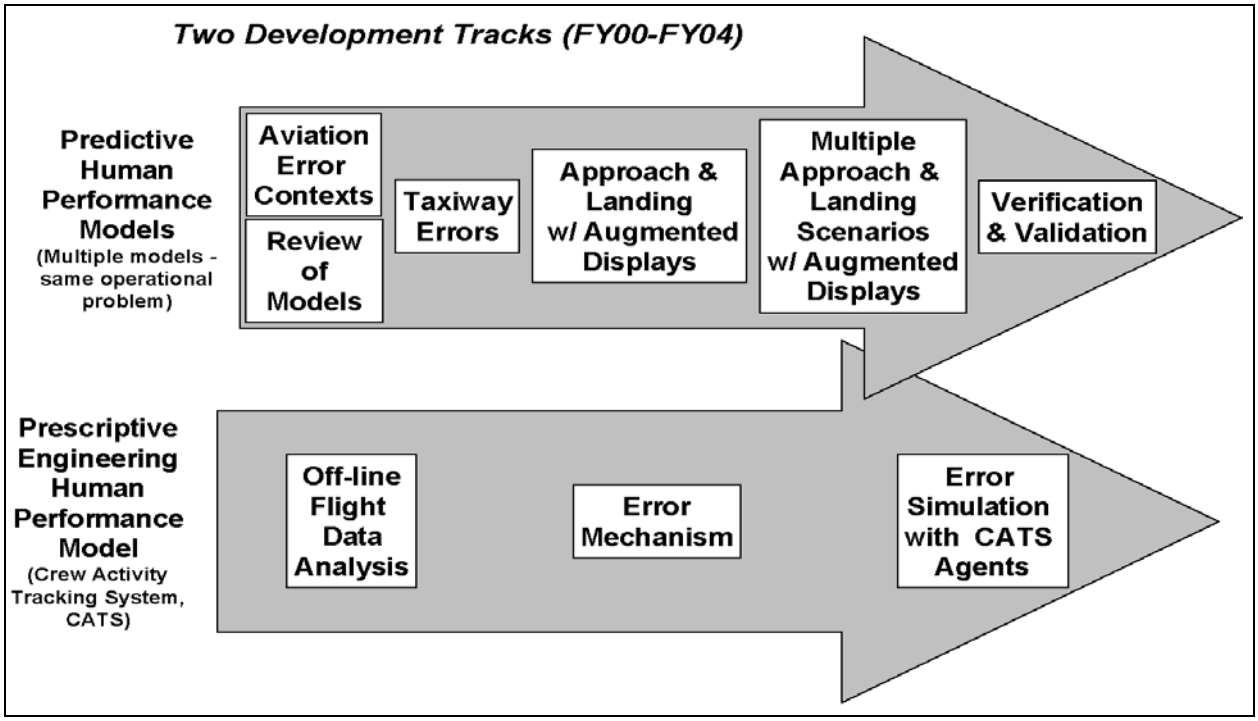

Figure E-1. Two model development tracks of the HPM element: Predictive Human Performance Models with multiple predictive models investigating a set of common problems; and, Prescriptive Engineering Human Performance Model.

Predictive Human Performance Models

From an initial review of past efforts in cognitive modeling, it was recognized that no single modeling architecture or framework had the scope to address the full range of interacting and competing factors driving human actions in dynamic, complex environments (Leiden, Laughery, Keller, French, Warwick & Wood, 2001). As a consequence, the HPM element sought to develop and extend multiple modeling efforts to further the current state of the art within a number of HPM tools. In 2001, five modeling frameworks were selected from a large group of responses to a proposal call for computational approaches for the investigation and prediction of operator behaviors associated with incidents and/or accidents in aviation. This was, in essence, a request for analytic techniques that employed cognitive modeling and simulation. The proposals were peer-reviewed with selection criteria including model theory, scope, maturity, and validation as well as the background and expertise of the respective research team.

All five of the predictive human performance modeling frameworks share common, important human characteristics. The models are:

(1) Generative -- Output results from the flow of internal model processes and is not "scripted";

(2) Have stochastic elements -- Simulation runs are not identical, even when all parameters defining the environment external to the human operator are held constant;

(3) Context sensitive -- Changes in the task-environment effect changes in simulation output.

Four of the five selected modeling frameworks were based on mature, validated, and integrative architectures which linked together embedded component processes of cognition with capabilities to construct representations of the task-environment and for simulations. The fifth modeling framework (A-SA) is a set of computational algorithms, more limited-in-scope, focused on attentional processes and the assessment of situational awareness that will be described below.

Additional characteristics of these five models are summarized in figure E-2.

The five predictive human performance models are:

ACT-R (Rice University; University of Illinois). *Atomic Components of Thought-Rational* is an experimentally grounded, open-source, low-level cognitive architecture developed at Carnegie Mellon University. ACT-R is based on the assumption that human cognition should be implemented in terms of neural-like computations on a very small time scale (50 ms–200 ms). A cognitive layer interacts with a perceptual-motor layer to create activation levels which determine both knowledge accessibility and goal-oriented conflict resolution.

Air MIDAS (San Jose State University). Air MIDAS is a version of *the Man-machine Integration Design and Analysis System* (MIDAS) developed as a joint Army-NASA program to explore computational representations of human-machine performance. Air MIDAS is driven by a set of user inputs specifying operator goals, procedures for achieving those goals, and declarative knowledge appropriate to a given simulation. These asserted knowledge structures interact with, and are moderated by, embedded models of cognition for managing resources, memory, and action.

A-SA (University of Illinois). *Attention-Situational Awareness* is a computational model developed at the University of Illinois. The underlying theoretical structure of the A-SA model is contained in two modules, one governing the allocation of attention to events and channels in the environment, and the second drawing an inference or understanding of the current and future state of the aircraft within that environment. Four factors are used to compute attention allocation within a dynamic environment; salience, effort, expectancy, and value. In turn, attentional allocation modulates situational awareness.

D-OMAR (BBN Technologies). The *Distributed Operator Model Architecture* was originally developed by BBN Technologies under sponsorship from the Air Force Research Laboratory. D-OMAR supports the notion of an agent whose actions are driven not only by actively seeking to achieve one or more goals, but also by reacting to the input and events of the world. It was designed to facilitate the modeling of human multi-tasking behaviors of team members interacting with complex equipment.

IMPRINT/ACT-R (Micro Analysis and Design, Inc.; Carnegie Mellon University; Army Research Laboratory). This hybrid framework integrates *Improved Performance Research Integration Tool (IMPRINT)*, a task network-based simulation tool developed by Micro Analysis and Design and *Atomic Components of Thought-Rational* (ACT-R), a low-level cognitive architecture developed at Carnegie Mellon University. This approach is meant to exploit the advantages of top-down control with the emergent aspects of bottom-up behavior for evaluating human performance in complex systems.

Model	*Type*	*Research Team*	*Demonstrated Sources of Pilot Error*
ACT-R	Low-level Cognitive with Statistical Environment Representation	Rice University University of Illinois	* Time pressure * Misplaced expectations * Memory retrieval problems
Air MIDAS	Integrative Multi-component Cognitive	San Jose State University	* Workload * Memory interference * Misperception
A-SA	Component Model of Attention & Situational Awareness	University of Illinois	* Misplaced attention * Lowered situation awareness
D-OMAR	Integrative Multi-component Cognitive	BBN Technologies	* Communications errors * Interruption & distraction * Misplaced expectation
IMPRINT / ACT-R	Hybrid: Task Network with Low-level Cognitive	Micro Analysis and Design, Inc. Carnegie Mellon University Army Research Laboratory	* Time pressure * Perceptual errors * Memory retrieval * Inadequate knowledge

Figure E-2. The five predictive human performance models, type of model, and demonstrated sources of pilot error.

Prescriptive Engineering Human Performance Model

CATS (San Jose State University/NASA Ames Research Center). The Crew Activity Tracking System (CATS) is a prescriptive engineering model which provides a representation of the task that the user is attempting to complete, a representation of how the task should be completed, and the capability to track and compare actual performance against prescribed performance. The model allows for error-detection and is being expanded to include mechanisms that produce observed operator errors.

HPM Element Approach and Scope

The 2003 HPM Conference and the resulting Conference Proceedings focused on the particular aviation safety-related problem of approach and landing with and without augmented displays. By plan, only the predictive human performance models addressed this problem. For this reason, the Prescriptive Engineering Human Performance Model, CATS, is not included in these proceedings. For more information, the reader is referred to articles on that topic (e.g., Callantine, 2002). The problem and approach described below refers only to the five predictive models of human performance.

The approach used in the AvSP HPM element involves applying different cognitive modeling frameworks to the analysis of a well-specified operational problem for which there is available empirical data of pilot performance in the task (see figure E-3). In 2001, the five different modeling frameworks were used to analyze a series of land-and-taxi-to-gate scenarios taken from a high-fidelity full mission simulation study that produced an extensive data-set of pilot performance. This completed 2001 effort is represented by the left-most panel of figure E-3. Overall, this approach enables the HPM Element to assess and contrast the predictive ability of a diverse range of human performance modeling frameworks while encouraging the advancement of these frameworks. For 2002-2003 (figure E-3 center panel), the five predictive modeling frameworks have been extended to the more complex problem of modeling pilot behaviors during approach and landing operations with and without the availability of a synthetic vision system (SVS)). This is in accord with the HPM Element's 2002 milestone objective requiring the development of cognitive models of an approach/ landing scenario with an augmented display. In 2003-2004 (figure E-3 right panel), these five models will focus on other specific approach and landing scenarios. A schematic representing the multiple off-nominal conditions (e.g., late runway reassignment; SVS display malfunction; and "go-arounds" because of cloud cover and runway traffic) to be investigated in the last years of the program is shown in the rightmost panel of figure E-3.

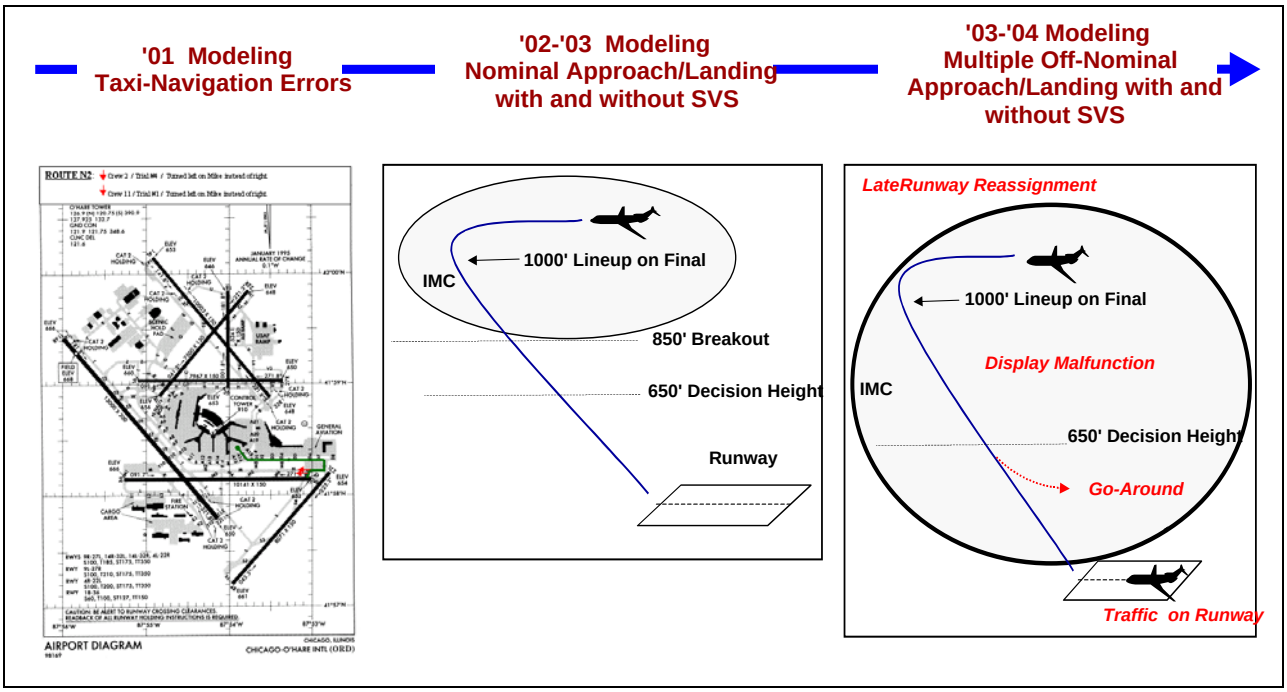

Figure E-3. Three phases of aviation safety-related problems addressed by the five Human Performance Modeling predictive models during 2001-2004.

Current HPM Efforts and Findings

In these conference proceedings, papers describing current accomplishments of the predictive human performance models of the HPM element are presented. The first two papers serve to set the stage for the modeling efforts which follow. The first paper in these proceedings describes a cognitive task analysis of the approach and landing phase of flight conducted by Keller, Leiden and Small. Next, is a discussion of a part-task human-in-the-loop simulation, the tested scenarios, and the data supplied to the five modeling teams by Goodman, Hooey, Foyle and Wilson. Following these two papers in these proceedings are descriptions of the modeling efforts and their results to date. Summaries of these five predictive human-performance-modeling efforts are given below.

In the first modeler's report in these proceedings, Byrne and Kirlik describe three central principles which guided their modeling approach: 1) the desire to create a dynamic, close-loop model of pilot cognition in interaction with the cockpit, aircraft, and environment; 2) the presumption that pilots are knowledgeable and adapted operators; and, 3) a focus on the allocation of visual attention as crucial to yielding important design and training-related insights. Their model, implemented in the ACT-R/PM cognitive architecture and referencing a statistical description of the environment, produced high-level predictions of gaze time that fit well with human-in-the-loop simulation data. Additionally, the model proved sensitive to the local properties of the SVS display, demonstrating that the type and format of presented flight symbology is a strong determinant of SVS usage. This suggests one line of focused investigation in which model predictions are used to assess a range of small variations to the symbology set of the SVS display in order to optimize pilot performance.

Next, in the report by Corker, Gore, Guneratne, Jadhav and Verma, the authors document their efforts to augment the standard Air MIDAS modeling architecture with an advanced vision model incorporating the affects of contrast legibility and visual search/reading time to better account for performance using a SVS display. To gain additional accuracy, the visual sampling model was calibrated and verified with an extensive, alternate empirical data set. The revised model generated predictions of pilot visual scanning behavior over three approach and landing scenarios. These model predictions explained 31% to 77% of the variance of the human-in-the-loop simulation data. Output from model simulations also permitted detailed inspection of the executed task sequences for both the pilot flying and pilot not flying. Analyses of these sequences indicate differences in task completion ordering, timing, and success between scenario conditions. This suggests possible vulnerabilities in crew coordination and timing resulting from specific situational demands. In another finding, the authors acknowledge that a better understanding of how flight crews select from redundant information sources is needed to improve fidelity.

Deutsch and Pew describe their efforts to implement a dedicated model of approach and landing within the D-OMAR simulation framework. The cognitive architecture evolved for this application focuses on multi-task behavior, the role of vision, and working memory. In simulation, the resultant models of the Captain, First Officer, and Air Traffic Controller working in concert demonstrated a commendable robustness by executing successful landings across five different scenarios circumstances. The model’s prediction that the availability of the SVS display would reduce time devoted to HSI display is matched in the human data. This finding supports the implication that information redundancy on the SVS display may reduce workload. The authors also note that additional scenario complexity can lead to better models by teasing out flaws. This was the case when certain D-OMAR model shortcomings only became apparent when a distracter aircraft was added to the scenario.

The paper by Lebiere, Archer, Schunk, Biefeld, Kelly and Allender details the unique integration of the low-level cognitive architecture, ACT-R, with the task network simulation tool, IMPRINT, to provide a viable approach for modeling complex domains. Functionality of the resulting model of approach and landing operations permitted sensitivity analyses of mission success rates to global parameters regarding latency of procedural, visual, motor, and auditory actions, as well as stochastic manipulations of decision-making times. These analyses provided important inferences regarding effective design objectives for both information display and procedures. Among other findings, the model found that pilot performance is very sensitive to the speed of visual shifts between widely separated information sources. Similarly, pilot performance proved highly sensitive to the overhead of communications with increases in the number and/or duration of communications acts rapidly deteriorating performance. Noteworthy in these modeling analyses was the apparent "performance tipping point" in which near-perfect mission success rates would suddenly plummet with only the slightest increase in parameter latency.

Wickens, McCarley and Thomas describe the modification of their algorithmic SEEV Model of attentional allocation in dynamic environments to the prediction of visual scanning during approach and landing operations. The refined algorithm for this application is based on the parameters of effort, expectancy, and value. This revised model, the Attention-Situation Awareness (A-SA) Model, accounted for roughly 30% - 80% of the variance in the scanning behavior seen in the human data. Surprisingly, the effort parameter added no predictive power to the model beyond

expectancy and value in this application. The authors do make a qualitative distinction between "good" and "poor" SA pilots based on latency to execute go-around maneuvers. They find that deviation from model predicted dwell times to the outside world clearly discriminated between these two categories of pilots. This is seen as supporting the model's ability to infer pilot SA. The authors also note that the large observed variance between individual pilot scanning behaviors (and resulting impact on model fit) may be attributable to one of two causes: 1) different pilot strategies of accessing information from redundant displays; or, 2) less-than optimal scanning behavior from some pilots. Again, it is asserted that the model can make that discrimination.

As will be seen in the following papers in these conference proceedings, the HPM predictive modeling efforts resulted in both design solutions and procedural recommendations to enhance the safety of SVS systems. The models identified potential problems that merit further investigation through human-in-the-loop simulations. Significant advancements to the state of human performance modeling were achieved by broadening the scope of the five models to include the aviation domain, and through the augmentation and expansion of specific modeling capabilities.

References

All these publications are available for download at the "publications link" at: http://humanfactors.arc.nasa.gov/ihi/hcsl/

Byrne, M.D. and Kirlik, A. (2003). Integrated modeling of cognition and the information environment: A closed-loop, ACT-R approach to modeling approach and landing with and without synthetic vision system (SVS) technology. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Callantine, T. (2002). Activity tracking for pilot error detection from flight data. NASA Contractor Report 2002-211406. Moffett Field, CA: NASA Ames Research Center.

Corker, K.M., Gore, B.F., Guneratne, E., Jadhav, A. and Verma, S. (2003). Human Performance Modeling Predictions in Reduced Visibility Operation with and without the use of Synthetic Vision System Operations. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Deutsch, S. & Pew, R. (2003). Modeling the NASA baseline and SVS-equipped approach and landing scenarios in D-OMAR. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Goodman, A., Hooey, B.L., Foyle, D.C. and Wilson, J.R. (2003). Characterizing visual performance during approach and landing with and without a synthetic vision display: A part-task study. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Keller, J., Leiden, K. and Small, R. (2003). Cognitive Task Analysis of Commercial Jet Aircraft Pilots during Instrument Approaches for Baseline and Synthetic Vision Displays. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Lebiere, C., Archer, R., Schunk, D., Biefeld, E., Kelly, T. and Allender, L. (2003). Using an integrated task network model with a cognitive architecture to assess the impact of technological aids on pilot performance. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.

Leiden, K., Keller, J. W., and French, J.W. (2001). Context of Human Error in Commercial Aviation. Contractor Report.

Leiden, K., Laughery, K.R., Keller, J. W., French, J.W., Warwick, W. and Wood, S.D. (2001). A Review of Human Performance Models for the Prediction of Human Error. Contractor Report.

Wickens, C., McCarley, J. and Thomas, L. (2003). Attention-Situation Awareness (A-SA) Model. In Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays, (David C. Foyle, Allen Goodman & Becky L. Hooey, Eds.). NASA Conference Proceedings NASA/CP-2003-212267.