NASA/CP—2003–212267

NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays

Edited by David C. Foyle, Allen Goodman, and Becky L. Hooey
Ames Research Center, Moffett Field, California

September 2003

The NASA STI Program Office ... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results ... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at ***http://www.sti.nasa.gov***

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Telephone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076-1320

NASA/CP—2003–212267

NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays

Edited by David C. Foyle
Ames Research Center, Moffett Field, California

Allen Goodman
San Jose State University, San Jose, California

Becky L. Hooey
Monterey Technologies, Inc., Santa Clara, California

Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays sponsored by the National Aeronautics and Space Administration and held at NASA Ames Research Center, Moffett Field, California March 6, 2003

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

September 2003

Available from:

NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320
301-621-0390

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
703-605-6000

Table of Contents

An Overview of the NASA Aviation Safety Program (AvSP) System-Wide Accident Prevention (SWAP) Human Performance Modeling (HPM) Element

David C. Foyle
NASA Ames Research Center
Moffett Field, California

Allen Goodman
San Jose State University
San Jose, California

Becky L. Hooey
Monterey Technologies, Inc.
Santa Clara, California

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 System-Wide 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 modeling and 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 which may not have been envisioned or tested.

When combined with nominal and off-nominal scenario human-in-the-loop testing, human performance modeling provides a complementary technique to develop systems and procedures that are tailored to the pilots' tasks, capabilities, and limitations. Because of its fast-time nature, 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, modeling and 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 permits the generation of very large sample sizes from which low-rate-of-occurrence events are more likely to be revealed. 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 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 track which consists of a single model of error detection (specifically a prescriptive engineering model of operator performance in context). The six models comprising the AvSP HPM element are listed and described below.

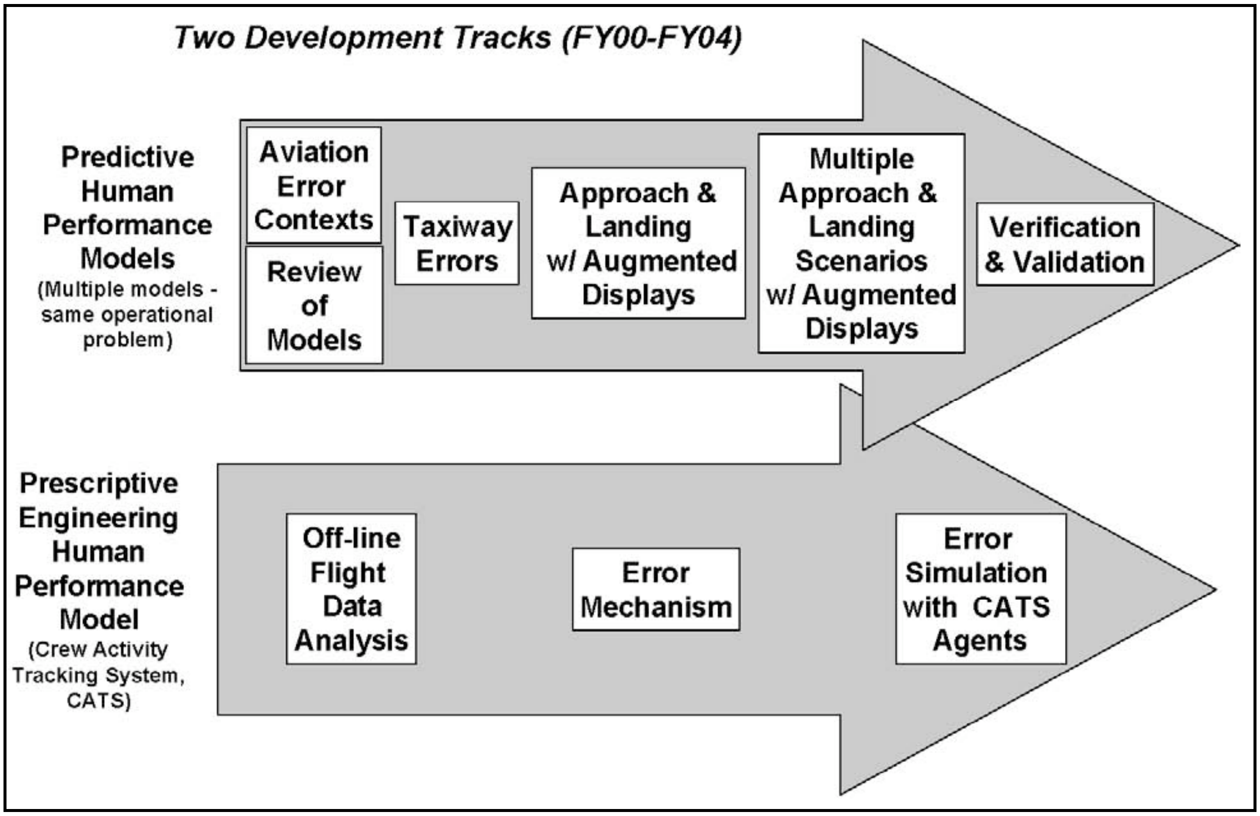


Figure 1. Two model development tracks of the HPM element: Predictive Human Performance Models (top) with multiple predictive models investigating a set of common problems; and, Prescriptive Engineering Human Performance Model (bottom).

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 extend 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 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 are held constant; and, are

(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 to run simulations. (The outstanding modeling framework is a more limited-in-scope set of computational algorithms 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 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 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 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 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 the modeling enterprise. For 2002-2003 (Figure 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 display. This is in accord with the HPM Element's 2002 milestone objective (MS 2.2.1/7) calling for the development of cognitive models of an approach/landing scenario with an augmented display. In 2003-2004 (Figure 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 3.

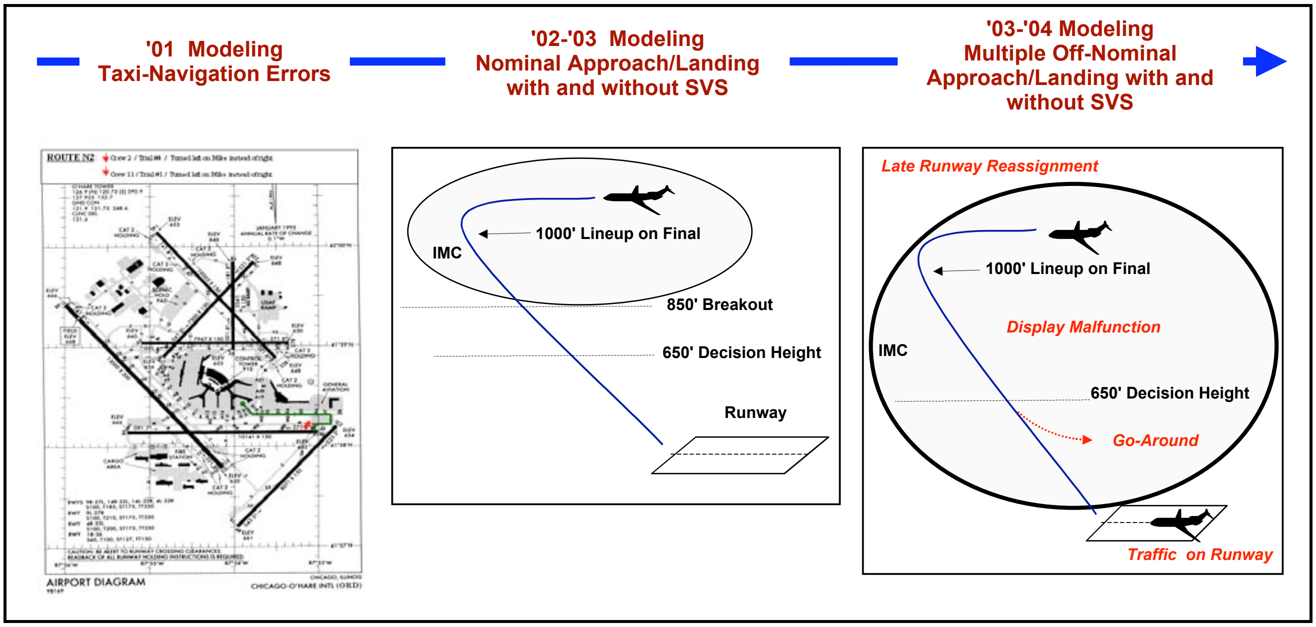


Figure 3. The aviation safety-related problems addressed by the five Human Performance Modeling predictive models during 2001-2004.

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.

AvSP SWAP HPM Bibliography

Listed below are the reports of the AvSP Human Performance Modeling element (as of November, 2003). Many of these reports are available for download from the NASA Ames Research Center Human-Centered Systems Laboratory (HCSL) website "publications" link:

http://human-factors.arc.nasa.gov/ihi/hcsl/

Predictive Human Performance Models

Byrne, M. D., & Kirlik, A. (in prep). Marrying cognitive and ecological analyses to support computational modeling of dynamic decision making in aviation. To appear in: A. Kirlik (Ed.), Working with Technology in Mind: Brunswikian Resources for Cognitive Science & Engineering. New York: Oxford University Press.

Byrne, M. D., & Kirlik, A. (2003, submitted). Using computational cognitive modeling to diagnose possible sources of aviation error. Manuscript submitted to The International Journal of Aviation Psychology.

Byrne, M. D., & Kirlik, A. (2002). Integrated Modeling of Cognition and the Information Environment: Closed-Loop, ACT-R Modeling of Aviation Taxi Errors and Performance. Technical Report AHFD-02-19/NASA-02-10, Institute of Aviation, University of Illinois at Urbana-Champaign.

Byrne, M. D., & 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. Technical Report AHFD-03-4/NASA-03-3, Institute of Aviation. University of Illinois at Urbana-Champaign.

Corker, K.M., Gore, B.F., Guneratne, E., Jadhav, A., & Verma, S. (2003). SJSU/NASA coordination of Air MIDAS safety development human error modeling: NASA aviation safety program. Integration of Air MIDAS human visual model requirement and validation of human performance model for assessment of safety risk reduction through the implementation of SVS technologies, (Interim Report and Deliverable NASA Contract Task Order #: NCC2-1307), Moffett Field, CA.

Deutsch, S., & Pew, R. (2001). Modeling human error in D-OMAR. BBN Report No. 8328. Contractor Report.

Deutsch, S. & Pew, R. (2002). Modeling human error in a real-world teamwork environment. Proceedings of the Twenty-fourth Annual Meeting of the Cognitive Science Society, 274-279, Fairfax, VA

Deutsch, S., & Pew, R. (2003). Modeling the NASA baseline and SVS-equipped approach and landing scenarios in D-OMAR. BBN Report No. 8364. Contractor Report.

Gore, B.F. (2002). Human performance cognitive-behavioral modeling: A benefit for occupational safety. In B. Chase & W. Karwowski (Eds.), International Journal of Occupational Safety and Ergonomics (JOSE), 8 (3), 339-351.

Gore, B. F. (2002). An emergent behavior model of complex human-system performance: An aviation surface related application. VDI Bericht 1675, 1 (1), 313-328, Düsseldorf, Germany: VDI Verl
Gore, B.F., & Corker, K.M. (2001). Human error modeling predictions: Increasing occupational safety using human performance modeling tools. In B. Das, W. Karwowski, P. Modelo, and M. Mattila (eds.), Computer-Aided Ergonomics and Safety (CAES) 2001 Conference Proceedings, July 28 - August 4, Maui, Hawaii.

Gore, B. F., and Corker, K. M. (2002). Increasing aviation safety using human performance modeling tools: An Air Man-machine Integration Design and Analysis System application. In M. J. Chinni (Ed). 2002 Military, Government and Aerospace Simulation, 34(3), 183-188. San Diego: Society for Modeling and Simulation International.

Gore, B.F., Verma, S., Jadhav, A., Delnegro, R., & Corker, K.M. (2002). Human error modeling predictions: Air MIDAS human performance modeling of T-NASA. NASA Ames Research Center Contract No.21-1307-2344. CY01 Final Report.

Goodman, A., Hooey, B. L., and Foyle, D. C. (2003). Developing Cognitive Models of Approach and Landing with Augmented Displays, NASA Milestone Report.

Keller, J. W., and Leiden, K. (2002). Information to Support the Human Performance Modeling of a B757 Flight Crew during Approach and Landing: RNAV. Contractor Report.

Keller, J. W., and Leiden, K. (2002). Information to Support the Human Performance Modeling of a B757 Flight Crew during Approach and Landing SVS Addendum. Contractor Report.

Lebiere, C., Biefeld, E., Archer, R., (2003) Cognitive models of approach and landing. Contractor Report.

Lebiere, C., Biefeld, E., Archer, R., Archer, S., Allender, L., and Kelley, T. D. (2002). Imprint/ACT-R: Integration of a task network modeling architecture with a cognitive architecture and its application to human error modeling. In M. J. Chinni (Ed). 2002 Military, Government and Aerospace Simulation, 34(3), 13-19. San Diego: Society for Modeling and Simulation International.

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

Leiden, K., Keller, J. W., and French, J., (2002). Information to Support the Human Performance Modeling of a B757 Flight Crew during Approach and Landing, 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.

McCarley, J. S., Wickens, C. D., Goh, J., and Horrey, W. J. (2002). A computational model of attention / situation awareness. Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society. 1669-1673. Santa Monica: Human Factors and Ergonomics Society.

Newman, R. L. (2002). Scenarios for "rare event" simulation and flight testing. Monterey Technologies Inc. / Crew Systems TR-02-07A.

Newman, R.L. and Foyle, D.C. (2003). Test scenarios for rare events. Proceedings of the Twelfth International Symposium on Aviation Psychology, 873-882. Dayton, OH: Wright State University.
Uhlarik, J. and Prey, C.M. (2002). Functional Allocation Issues and Tradeoffs (FAIT) Analysis of Synthetic Vision Systems (SVS). Contractor Report.

Wickens, C. D. (2002). Spatial Awareness Biases (ARL-02-6/NASA-02-4). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., & McCarley, J. S. (2001). Attention-Situation Awareness (A-SA) Model of Pilot Error (Final Technical Report ARL-01-13/NASA-01-6). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., McCarley, J. S. and Thomas, L. (2003). Attention-Situation Awareness (A-SA) Model, Contractor Report.

Prescriptive Engineering Human Performance Model

Callantine, T. (2001). Agents for analysis and design of complex systems. Proceedings of the 2001 IEEE International Conference on Systems, Man, and Cybernetics, 567-573.

Callantine, T. (2001). Analysis of flight operational quality assurance data using model-based activity tracking. SAE Technical Paper 2001-01-2640. Warrendale, PA: SAE International.

Callantine, T. (2001). The crew activity tracking system: Leveraging flight data for aiding, training, and analysis. Proceedings of the 20th Digital Avionics Systems Conference, 5.C.3-1-5.C.3-12 (CD-ROM).

Callantine, T. (2002). A representation of air traffic control clearance constraints for intelligent agents. In A. El Kamel, K. Mellouli, and P. Bourne (Eds.), Proceedings of the 2002 IEEE International Conference on Systems, Man, and Cybernetics, #WA1C2, (CD-ROM).

Callantine, T. (2002). Activity tracking for pilot error detection from flight data. Proceedings of the 21st European Annual Conference on Human Decision Making and Control, Glasgow, 16-26.

Callantine, T. (2002). CATS-based agents that err. NASA Contractor Report 2002-211858. Moffett Field, CA: NASA Ames Research Center.

Callantine, T. (2002). CATS-based air traffic controller agents. NASA Contractor Report 2002-211856. Moffett Field, CA: NASA Ames Research Center.

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

Cognitive Task Analysis of Commercial Jet Aircraft Pilots during Instrument Approaches for Baseline and Synthetic Vision Displays

John Keller, Kenneth Leiden, and Ronald Small
Micro Analysis & Design, Inc.
Boulder, Colorado

Abstract

This paper describes flight crew tasks and procedures during commercial jet instrument approaches. The objective is to provide the necessary background information to support computational human performance modeling under two experimental conditions: (1) a baseline that assumes current aircraft displays, controls and navigation guidance (e.g., a Boeing 757 flying instrument landing system (ILS) and area navigation (RNAV) approaches); and, (2) the addition of a prototype synthetic vision system (SVS).

1. Introduction

Human performance modeling (HPM)[1] is a valuable research tool for understanding new systems and their impact on human task performance and workload, without resorting to the more costly methods of human-in-the-loop experiments or simulations. This paper describes task analyses of jet transport pilot tasks during approach to landing, using instrument landing system (ILS), area navigation (RNAV), and synthetic vision system (SVS) display methods in support of on-going HPM research.

The ultimate goal of the research is to understand the impact of SVS on pilot performance, workload, and other human factors considerations. The ILS and RNAV models will serve as baseline conditions. HPM and simulation can be used to compare the baseline and SVS models to illustrate advantages and disadvantages in pilot performance.

Using HPM for understanding the benefits and drawbacks of any new system must be part of a continuum of research methods. The continuum ranges from pencil and paper (mathematical) analyses, to pre-development modeling and human-in-the-loop simulations, all the way to usability testing with actual systems or components. Good practice dictates starting with the less expensive evaluation methods (e.g., mathematical analysis) and progressing to user tests before fielding a new system. Typically, the less expensive methods help focus the objectives of the more expensive methods, thus maximizing the application of system engineering resources (Small & Bass, 2000).

1.1. Context for HPM Task Analyses

The NASA Aviation Safety Program (AvSP) was created to perform research and develop technology to reduce the rate of fatal aircraft accidents in the US. Under AvSP, the System-Wide Accident Prevention (SWAP) project uses current knowledge about human cognition to develop mitigation strategies to address current trends in aviation accident and incident profiles. System-Wide Accident Prevention is comprised of four elements, one being Human Performance Modeling. The objective of the HPM Element is to develop predictive capabilities to identify likely performance improvements or error vulnerabilities during system operations. During the time period of interest (FY02), this element investigated the application of HPM to predict the human performance of flight crews using SVS in the cockpit. It began with analyses of pilot tasks to support HPM development. SVS is a developmental system whose details have not been fully determined. It is intended to present to the pilots a clear, 3D, out-the-cockpit view of terrain, obstacles, traffic, and runways, <u>regardless</u> of the actual visibility or weather conditions.

1.2. Baseline & SVS Models

The first step in the FY02 SVS effort was to analyze pilot tasks in order to create a baseline human performance model of the flight crew *without* SVS. In other words, the baseline model will represent today's flight deck equipment and operations. NASA decided that the example flight deck for this HPM effort would be the Boeing 757 (B757). This decision was driven by the fact that preliminary

---

[1] All acronyms are defined in Section 9 at the end of this paper.

SVS flight tests used a NASA-owned B757. Data collected from the flight tests may be used for comparison with the HPM predictions. The NASA HPM element directed the FY02 effort to focus on the approach and landing phases of flight because improved pilot situational awareness of terrain and obstacles during approach and landing is expected to be one of the biggest benefits of SVS.

Micro Analysis and Design (MA&D) began the *baseline* (no SVS) effort by analyzing B757 approach tasks for ILS- and RNAV-guided approaches. Then, MA&D conducted a preliminary task analysis for the SVS HPM.

2. Technical Objectives

The objectives of the research reported herein are to:

- Analyze and enumerate the pilot tasks for the baseline conditions (i.e., B757 approaches using ILS and RNAV guidance).
- Analyze and describe the pilot tasks for the SVS condition.

Assumptions for the analyses include:

- Analytic results from above will be used by human performance modelers, so enough context is needed to model a rich enough environment for valid HPM experiments.
- HPM modelers do not have aviation domain expertise; therefore the task analyses must include all relevant task information (i.e., the *what*, *how*, and *why* for each task).
- Pilot tasks include visual, auditory, cognitive, and psychomotor elements.
- Approaches do not include abnormal or emergency situations; although, since safety is paramount, planning for a missed approach or go-around is always part of flying an approach.

3. Background Information

Although this research assumes a B757 aircraft, this background information applies to practically all commercial jet aircraft. Air carriers (i.e., scheduled and charter airlines, cargo carriers, and business jets) typically file instrument flight rules (IFR) flight plans. Aircraft on IFR flight plans follow air traffic control (ATC) instructions; in return, ATC assumes responsibility for the safe separation of IFR aircraft.

Because of the complexity of the approach phase and the potential benefit offered by SVS, the initial HPM research focuses on this phase of flight. Figure 1 shows the relationship of the approach phase to the other phases of a *normal* flight (adapted from Alter & Regal, 1992).

While the missed approach and subsequent divert phases are shown in Figure 1, their occurrence is rare among professional pilots. Two of the professional pilots with whom we spoke estimated the occurrence of actual missed approaches to be about one missed approach per year per pilot. Therefore, based on an average of 20 landings per month per pilot, there would be 1 missed approach per 240 landings. Even though rare, a missed approach is still considered a normal phase of flight because pilots prepare for it each time they fly an approach. However, abnormal or emergency situations are presently outside the scope of this research.

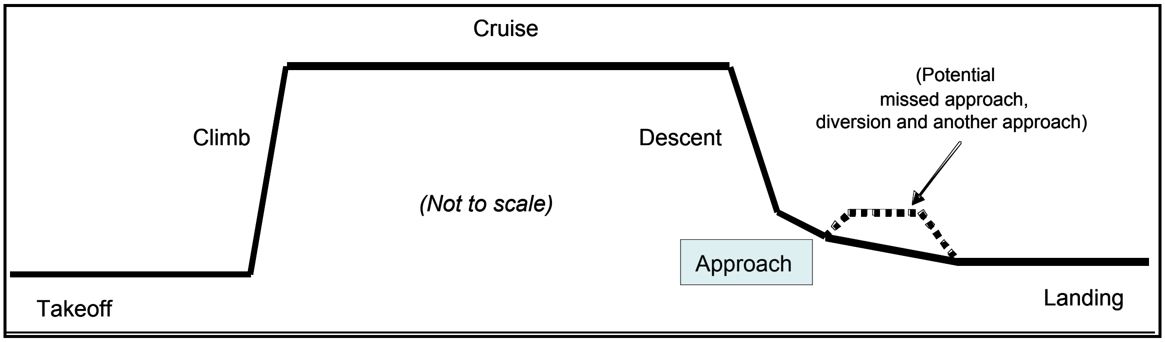


Figure 1. Normal phases of flight (approach phase highlighted).

The following sub-sections describe the approach phase and the four types of landing guidance that are relevant to our objectives: visual, ILS, RNAV, and SVS.

3.1. Approach Phase of Flight

The approach phase, which is the focus of the task analyses, begins at the bottom of descent and ends at main wheel touchdown in the landing phase. The purpose of an approach is to transition the aircraft in a carefully prescribed manner from typical intermediate altitudes after descent to a position, speed and configuration from which the pilots can land their airplane. During the approach, pilots normally follow published approach procedures, which designate mandatory courses, altitudes, and oftentimes speeds to a particular runway. Published approaches safely and expeditiously guide arriving aircraft into an airport by keeping aircraft away from high terrain, obstacles (e.g., radio antennas), aircraft departing from the airport, the approaches to other runways, and traffic patterns from nearby airports. Published approaches are most useful when visibility is poor because they enable the pilots to find the runway when they would not be able to otherwise. ATC gives permission, or *clearance*, to fly a specific approach, which is normally based upon weather, wind, and traffic conditions.

3.2. Approach Types

There are two basic types of approaches: visual and instrument. Instrument approaches are further delineated into precision and non-precision approaches. A non-precision approach provides only lateral guidance to the runway, whereas a precision approach provides both lateral and vertical guidance. Lateral guidance enables the pilots to align their aircraft with the runway centerline before landing. Vertical guidance enables the pilots to fly a steady descent angle, known as a *glide path,*[2] so as to land within approximately the first 1000 feet of the runway.

This HPM research focuses on two existing types of precision approaches (ILS and RNAV), as well as an “under development” SVS approach, which is, in essence, a combination of visual and precision approach types. Following, we explain a visual approach, and then an instrument approach with emphasis on a precision approach (ignoring the non-precision approach type for the remainder of this paper).

3.2.1. Visual Approach

A visual approach is when the pilots can see the airport and runway from a distance (usually at least 10-20 miles away). In this situation, pilots commonly follow other aircraft in the landing sequence

[2] The term glide path is a misnomer because the aircraft does not glide to landing. Rather, the descent to landing is a powered maneuver flown within specific tolerances of airspeed, thrust, configuration, pitch and roll.

and are responsible for spacing from the other aircraft, avoiding terrain and obstacles, and aligning their aircraft with the runway. Courses and altitudes are the pilots' choice, as long as they remain in the landing sequence. A typical ATC clearance for a visual approach is, "*NASA-113, Santa Barbara Approach: Follow the 737 ahead; cleared for a visual approach to Runway 7. Contact the tower for landing clearance*." This clearance identifies the aircraft (NASA-113) to whom ATC is talking (since the radio is a "party line" communication method), the ATC facility who is calling (Santa Barbara Approach), the current instructions ("follow the 737 ahead"), the clearance ("cleared for a visual approach to Runway 7"), and directions for what to do next (contact the tower).

Runway numbers are determined by the magnetic heading of an aircraft as it aligns with the runway centerline. So, an aircraft aligned with Santa Barbara's Runway 7 has an approximate heading of 070°. An airplane aligned with the same stretch of runway pavement, but at the opposite end and facing the opposite direction, has a heading of 250°, which is the reciprocal of 070°. That runway's number is 25. The single degree units are dropped, as is a leading zero, in most cases. Aviation compasses have 000° or 360° as due north, 090° as east, 180° is south, and 270° is west. (A runway facing north is numbered 36 not 0.)

All major airports have runways with specialized lighting systems, which are most useful at night or when visibility is limited due to weather (e.g., fog). There is lighting that outlines the runway and some that extends from the runway threshold along the approach path (up to 3000 feet in length) to help visually guide the pilots to the runway threshold. Most of the lights along the approach path and runway are designed to provide <u>lateral</u> guidance to the pilots. However, another type of lighting is the visual approach slope indicator system (VASI). The VASI provides pilots with <u>vertical</u> guidance (i.e., glide path) information by projecting narrow beams of light along the glide path. The lights are situated such that the pilots see white lights when above the glide path, and red lights when below. When the aircraft is on glide path, the VASI light bars show red over white (Figure 2). VASI lights are visible as far as 20 miles away in clear weather.

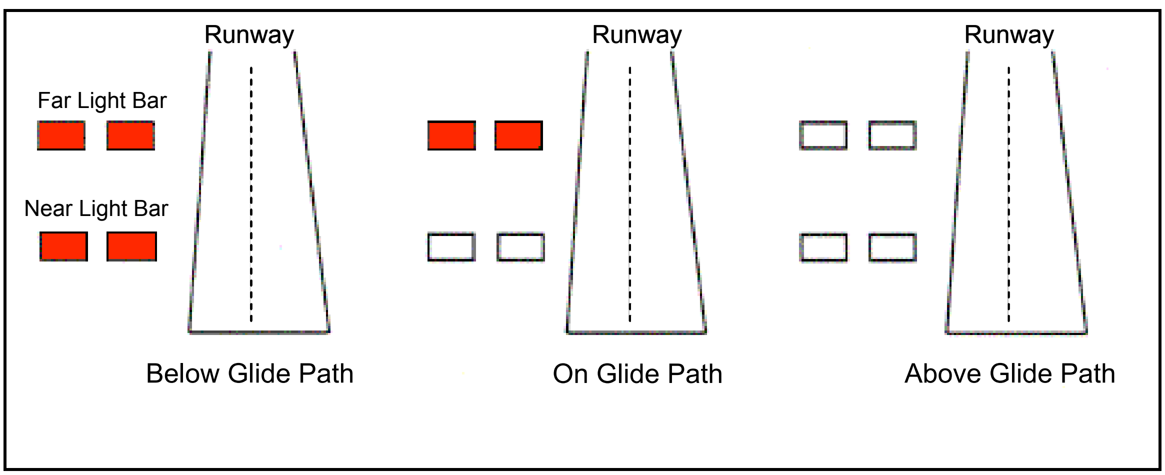


Figure 2. VASI lights and their meaning.

Even when visibility is good, pilots study, have available, and use applicable portions of the published instrument approach procedure for the designated runway as a prudent safety measure. Also, published procedures contain such useful information as communication radio frequencies, navigation radio frequencies, a diagram of the airport and runway with obstacle positions, and recommended courses and altitudes (which are especially useful at unfamiliar airports).

3.2.2. Instrument Approach

A typical ATC clearance for an instrument approach is "*NASA-113, Santa Barbara Approach: Fly heading 120; cleared for ILS Runway 7; maintain 3000 and 200 knots until established on the approach. Contact the tower at ROCKY*." The similarities to the visual approach clearance should be apparent. The differences are in the level of control exercised by ATC, usually because pilots cannot see the airport or surrounding terrain and obstacles. In this example, Santa Barbara Approach needs NASA-113 to maintain a specific heading (120°) to intercept the lateral approach guidance to the runway. Also, ATC requires a specific altitude ("maintain 3000" – feet are implied) and speed ("200 knots") until the aircraft is in a position to safely descend toward the specified runway on the published approach procedure (the ILS to Runway 7 – Santa Barbara is implied). The last portion of the clearance tells NASA-113 to radio the tower when at a navigation location called ROCKY. In aviation terminology, ROCKY is known as a *fix*. In this hypothetical example, ROCKY is a specific type of fix – the *final approach fix* for Runway 7 at Santa Barbara Airport.

As implied by the above instrument approach clearance explanation, published approaches are comprised of segments: initial, intermediate, and final. Figure 3 illustrates these segments using two views: the top portion shows the altitude or vertical profile; the bottom portion shows the course or lateral view. The initial approach segment is for beginning to descend toward the final approach fix altitude, and for beginning to align with the designated runway. For example, if the aircraft flew from New York to California (east to west) and if the destination airport, say Santa Barbara, is landing to the northeast on Runway 7 (due to its current winds), the pilots need to fly from a southwesterly heading to an northeasterly heading to eventually align with the runway. The initial approach segment starts this process. The intermediate segment continues the alignment and descent process, and usually includes slowing and configuring (i.e., lowering the landing gear and flaps).

So, by the time the aircraft reaches the final approach fix (FAF), all that remains for the pilots to do is descend at an appropriate rate and speed while remaining aligned with the runway centerline in the proper landing configuration. The final segment is from the FAF to the landing, or to the missed approach, if the pilots decide to abandon the landing. The overall purposes of the initial and intermediate segments are to ensure the aircraft is in the correct position, configuration, and orientation relative to the runway, at the FAF.

When the final approach segment proceeds as desired, it is known as a "stable" approach. A typical FAF altitude is about 1500 feet above the runway altitude. As the aircraft descends from the FAF to the landing, there are imaginary *gates* at 1000 feet above ground level (AGL) and 500 feet AGL. Most air carriers require their pilots to be stable at one or both of these gates (even for visual approaches). A stable approach is when the aircraft is in the correct configuration, and within established tolerances for speed, descent rate, attitude, and engine thrust. The aircraft must also be within tolerances for the published course and glide path. If any one of these conditions is not met, executing a missed approach is the appropriate action.

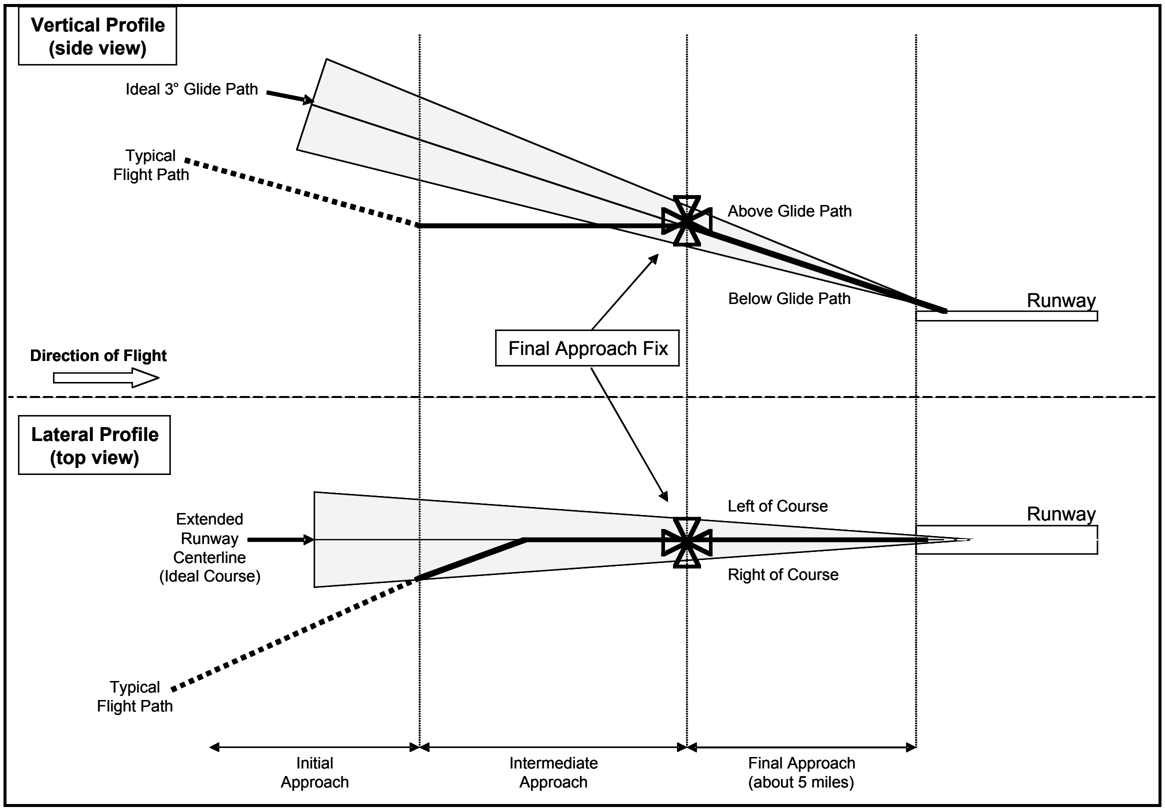


Figure 3. Two views of a typical instrument approach.

3.2.3. ILS Approach

An instrument landing system (ILS) approach is one type of precision approach. The system is comprised of three transmitters near the runway, and a receiver in the cockpit that the pilots tune to the appropriate frequency for the specific runway. The **localizer** transmitter provides lateral guidance aligned with the runway centerline. The **glide slope** transmitter provides vertical guidance to direct the aircraft along a glide path (typically 3°) that will intersect with the runway about 1000 feet from the approach end or threshold. Although 3° is a typical glide slope, false glide slopes, due to reflections of the signal off the ground, can occur at around 9°. To avoid flying a false glide slope, ILS approach procedures require that aircraft intercept the glide slope at an altitude low enough to inhibit false glide slope capture, but high enough to avoid terrain and obstacles (shown in the top half of Figure 3 by the typical flight path intercepting the glide slope cone from below).

Marker beacons signal when the aircraft flies over a particular spot on the ground. The beacon transmits vertically, so that the combination of the three ILS transmitters gives the pilots a definite location in 3D space. For ILS approaches, there is a marker beacon at the FAF; this marker beacon is also known as the *outer marker*. The altitude at which a vertical line from the *outer marker* intersects the ILS glide slope is noted on published ILS approach charts. Thus, if the aircraft is aligned with the ILS glide slope, it should cross the outer marker at the specified approach chart altitude. If the actual altitude differs from the charted altitude by a significant amount, the pilots should suspect a false glide slope and take appropriate actions (e.g., execute a missed approach if the runway is not in sight). A light on the ILS receiver unit in the cockpit flashes when the aircraft passes over a marker beacon. The *middle marker* is usually about 3000 feet from the runway threshold.

There are several *categories* of ILS based upon system accuracy, which help the pilots fly closer to the ground before reaching **decision height** – the minimum height above the runway at which the pilots decide to land or execute a missed approach. We are only concerned with Category I, the least accurate, with a typical decision height of 200 feet above the runway. If the aircraft is on the glide slope, crossing the middle marker occurs simultaneously with reaching the 200-foot AGL decision height.

As with generic instrument approaches, ILSs have three segments. The initial segment begins when ATC issues a clearance (altitude, heading, and possibly speed) that will result in the aircraft intercepting the localizer. Normally, only a single clearance is needed. However, when the controller must slow multiple aircraft for spacing and sequencing or when the instrument approach procedure requires it, this segment may necessitate multiple clearances (for any combination of altitude, heading, and speed) for each aircraft. In any case, the last of the clearances in this segment will place the aircraft on a heading to intercept the localizer, and at an altitude to intercept the glide slope from below. To simplify the scope of this work, the assumption is made that this segment will involve a single clearance (as noted in the example at the start of Section 3.2.2).

Once the aircraft has intercepted the localizer, the intermediate approach segment begins. The pilots fly the localizer inbound and continue following ATC instructions for altitude and speed, if given. The last altitude assigned by ATC is often the glide slope intercept altitude (GSIA), and is usually noted on the approach chart. However, if ATC directs an altitude that differs from the altitude specified on the approach chart, then the ATC-directed altitude takes precedence over the altitude from the approach chart (as is the case with any discrepancy between ATC clearances and an approach chart). The intermediate approach segment ends, and the final approach segment begins, when the aircraft intercepts the glide slope. This point is usually also the FAF, if the aircraft is at the GSIA. If ATC directs a different altitude than the published GSIA, the final approach still begins at the actual glide slope intercept, but will result in a shorter or longer final approach segment – shorter if the actual GSIA is below the published GSIA, longer if above.

During final approach, the aircraft descends along the ILS glide slope while maintaining alignment with the runway centerline via the localizer. No later than decision height, the pilots must be able to see the runway environment (i.e., the runway itself or any approach lights) to proceed with the landing. If the pilots cannot see the runway environment, or for any reason decide that it is imprudent to land (e.g., due to an unstable approach or wind shear), they execute a missed approach. The final approach segment ends upon landing or if executing a missed approach. A missed approach is a portion of the published approach that directs the aircraft to a safe altitude and location so that the pilots may attempt another approach, or divert to a different airport. Pilots begin a missed approach by accelerating, climbing, and raising the landing gear and, usually, the flaps. Then they typically inform ATC of their missed approach and follow the charted missed approach instructions (if ATC issues no other instructions). Because actual missed approaches are rare, the later task analyses (Section 4) only address planning for a missed approach, not flying one.

3.2.4. RNAV Approach

Virtually all modern commercial aircraft use an Area Navigation (RNAV) system for flight navigation. An RNAV system is comprised of several independent navigation subsystems, both internal and external to the aircraft, as well as subsystems to program and update the desired 3D route of flight. The RNAV system combines information from the navigation subsystems as a means to crosscheck and verify the computed location to the required level of accuracy. Based upon the actual computed aircraft position and the planned route of flight (entered by the pilots before flight and updated as needed during flight), the RNAV system provides guidance signals which the pilots or autopilot can follow to capture and fly the desired 3D path.

The RNAV flight path is a series of waypoints between takeoff and landing. A waypoint may be a navigation fix, such as an FAF, and usually has the following properties: latitude, longitude, inbound course, outbound course, crossing altitude, speed, and a unique name (such as ROCKY in the example in Section 3.2.2). Paths between waypoints are analogous to the approach segments, described earlier, in that waypoints delineate flight path changes in the form of changes to course, altitude, speed, or a combination of these parameters. The RNAV system's objective is to guide the aircraft to the next (*active*) waypoint.

A naïve observer might not see any difference between flying an RNAV approach and an ILS approach. Both approach types use waypoints to separate segments of the approach and both use a published decision height or altitude to define the point by which the pilots must decide to land or to execute a missed approach. The main difference is that the RNAV approach does not use ground-based localizer, glide slope, or marker beacon signals. Instead, an RNAV approach uses the same computations as during the rest of the flight. As such, RNAV is less accurate, and so RNAV *decision altitudes* are higher than ILS *decision heights*.

Another major difference involves the behavior of the aircraft under autopilot control. For an ILS, the autopilot will fly the glide slope all the way to the ground, unless the pilots take control. But, due to the inherent inaccuracies of RNAV, the autopilot will automatically execute a missed approach at the decision altitude. There are also minor terminology differences: RNAV uses lateral navigation (LNAV) for course guidance, not a localizer; vertical navigation (VNAV) for altitude and altitude change guidance, not a glide slope; and, has a decision altitude, not a decision height.

Figure 4 shows an example of a pilot's RNAV approach chart used in NASA simulations (i.e., not operationally certified). While much of the detail is outside the scope of this discussion, the following paragraphs explain some of the items.

Upper right corner:

- Title of the approach – in this case the RNAV (GPS) RWY 33L, Santa Barbara Muni. RNAV signifies the approach type. GPS refers to the type of navigation system used to determine positions. (GPS is global positioning system, a network of satellites used for determining the latitude, longitude and altitude of vehicles that have GPS receivers.) RWY 33L is an abbreviation for Runway 33 Left. The runway is at the municipal airport in Santa Barbara, whose three-letter identifier is SBA.

Upper center section (in boxes):

- Various information and radio frequencies.

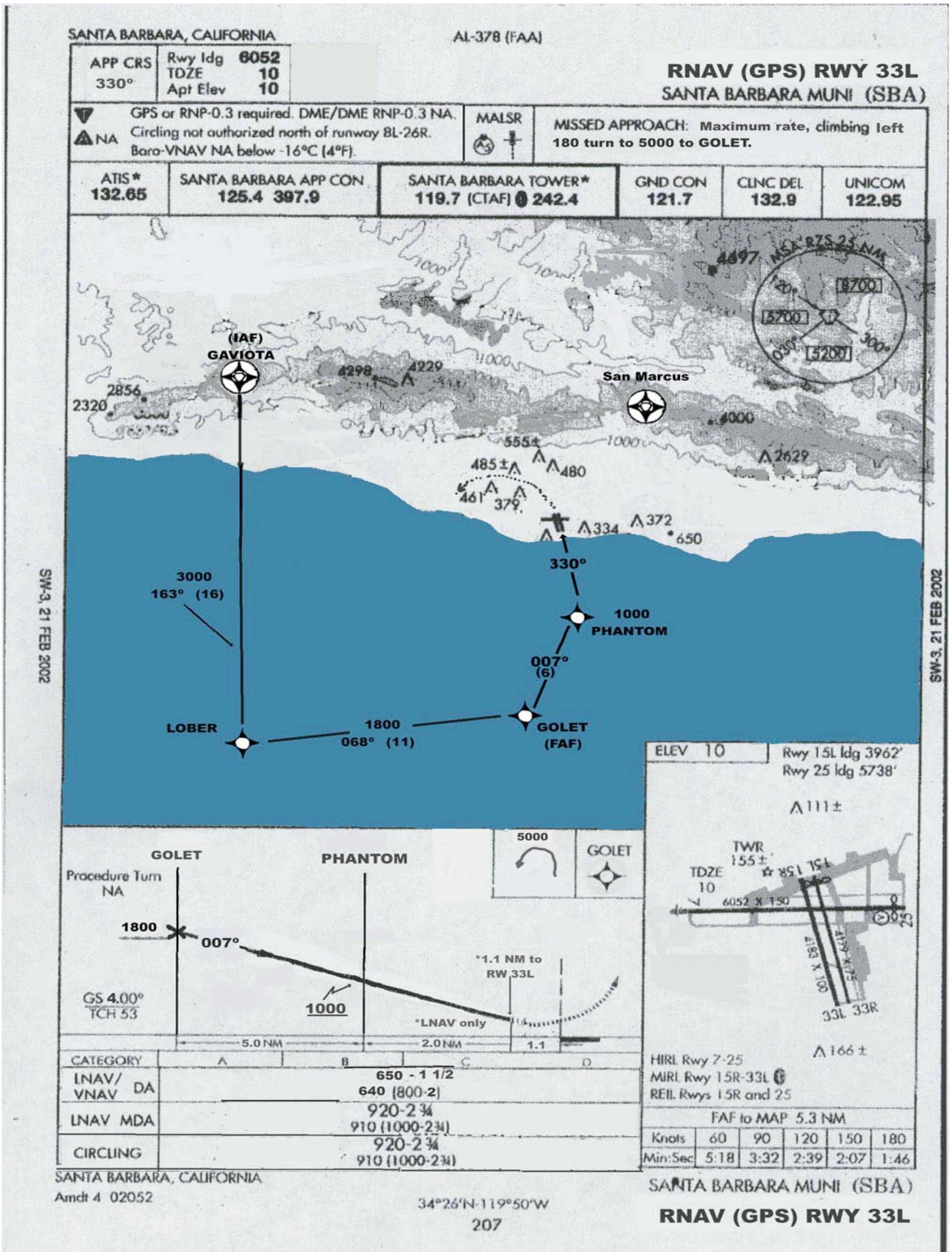


Figure 4. RNAV approach chart for Santa Barbara Airport’s Runway33L.

Middle section:

- A graphic that shows the relevant geography around the Santa Barbara airport, and the approach's route of flight superimposed on the geography image. The image is oriented north up, with the coastline clearly discernable.
- Terrain elevation lines depict the high terrain near the airport. Dots with numbers give the location (dot) and altitude (in feet above mean sea level (MSL)) of mountains, hills, and ridges.
- Upside-down Vs with dots and numbers show the location (dot) of obstacles, usually radio towers or buildings, and their height in feet MSL.
- Black stars with white middles are waypoints. Waypoint names are unique; in this example they are GAVIOTA, LOBER, GOLET, and PHANTOM. The GAVIOTA waypoint is the initial approach fix (IAF), which starts the approach. GOLET is the FAF.
- The approach's initial segment is from GAVIOTA to LOBER and has a course line of 163° at 3000 feet MSL for 16 nautical miles (nm).
- The intermediate segment is from LOBER to GOLET at 1800 feet, 068° and 11 nm.
- The final segment is from GOLET to the runway and includes PHANTOM. GOLET is offset from the runway centerline, probably because traffic for Runway 33R (Runway 33 Right) comes in from the other direction, and since RNAV is not as precise as ILS guidance, there is a dog-leg to the final runway centerline alignment.
- Runway graphic – shows the runway configuration and relative alignment. (A more detailed graphic of the runways is in the lower right corner of the chart.)
- Dashed line with arrow – depicts the missed approach path, which is also described in words right below the approach chart's title.

Lower right section:

- Runway graphic and details – shows the runways in black and taxiways in gray.
- Numbers alongside each runway give its dimensions (length x width).
- Numbers at the end of each runway give its name (7, 33L, 33R, 25, 15L, 15R).
- Symbols at the ends of some runways denote the lighting system for that runway.
- Upside-down Vs with dots and numbers show obstacle positions and heights.
- In the upper left corner is the field elevation, 10 feet MSL, which is the height of the highest terrain in the runway, taxiway and parking areas of the airport.

Lower left section:

- This graphic illustrates the vertical approach path from GOLET, the FAF, to the runway, including the missed approach (dashed line and arrow).
- Below the vertical profile are the various approach *minima*, the cloud ceiling and visibility criteria for flying this approach. The first row shows the LNAV/VNAV decision altitude (DA) of 650 feet MSL, with minimum visibility of 1.5 miles. Beneath those numbers are 640 for feet AGL (which is the 650 feet MSL DA minus 10 feet field elevation), and two numbers in parentheses, "800-2". The 800 is the lowest cloud ceiling height in feet MSL that the airport weather observers can give for pilots to legally begin the approach. The 2 is the lowest weather observation visibility for pilots to begin the approach. Both the observed ceiling and visibility must be above minima (800 and 2) for pilots to legally begin this published approach.

3.2.5. SVS Approach

Because the synthetic vision system (SVS) is under development and evolving, the information herein should only be considered a "best estimate" as of this writing. With that caveat in mind, the overall concept and anticipated use of SVS is to help pilots operate in poor visibility conditions (Norman, 2001). Therefore, an SVS approach is similar to either an ILS or RNAV approach, except that the pilots look at the SVS display, rather than out the window.

SVS is a cockpit system that allows pilots to view a computer-generated image of what is ahead of the aircraft, independent of weather or time of day (NASA Langley, 2001). The system will use satellite navigation signals to orient the presented image with the actual location of the aircraft relative to the terrain. This allows the terrain image to dynamically represent what the pilots would see out the front windows on a clear day as the flight progresses.

Because pilots are concerned with more than just terrain information, SVS is a very complex system with components in three major categories: sensors and database, computation, and display. Components in the sensors and database category include: forward looking infrared, millimeter wave radar, weather radar, navigation database, aircraft state data, and hazard information systems. The components in the computation category include: a dedicated computer for performing perspective transformations, data fusion, image object detection, display generation, integrity monitoring, and interface communication; and other existing aircraft subsystems, such as the RNAV, ground proximity warning, and central alert and warning systems. The components in the display category are: the primary flight display, navigation display, vertical situation display, head-up display, electronic moving map and other auxiliary displays (Norman, 2002).

3.2.6. Approach Summary

We have described four specific types of guidance for aircraft approaches: visual, ILS, RNAV, and SVS. The choice of which approach type to actually fly for any given situation depends upon the visibility, the type of guidance equipment available for the designated runway, and the corresponding instrumentation available on the flight deck.

The approach types are similar in that the ultimate objective is for the pilots to land on the designated runway. However, they are also very different in that a visual approach primarily relies on a pilot's vision, experience, and judgment. An ILS approach relies on ground-based transmitters. RNAV mainly uses internal systems for guidance. And, SVS has the goal of combining the best features from the previous three types. Because SVS is primarily a display system from the pilot's perspective, we next describe the baseline aircraft displays and controls and then discuss how the SVS enhances the baseline displays.

3.3. Pilot Displays & Controls

The following explanations of pilot displays and controls are specific to the Boeing 757 aircraft, but virtually all modern jet aircraft have analogous displays and controls, with similar appearances, modes, and control capabilities. First, we describe displays and controls that are common to all types of approaches; then we describe displays and controls unique to the ILS, RNAV, and, most importantly, SVS approaches.

3.3.1. Common Displays & Controls

The B757 flight deck is referred to as a "glass cockpit" because computers present aircraft attitude, navigation, and system information in an integrated and graphical fashion on flight-worthy computer displays. The left half of Figure 5 shows the B757 flight deck. The right half shows an older B727 flight deck that has dedicated gages for individual pieces of information. It was the change from many independent gages to combined information on fewer computer screens that led to the phrase "glass cockpit." For example, there are 15 engine gages in the center section of the front instrument panel in the B727, but one computer screen for similar information in the B757 (red circles in Figure 5).

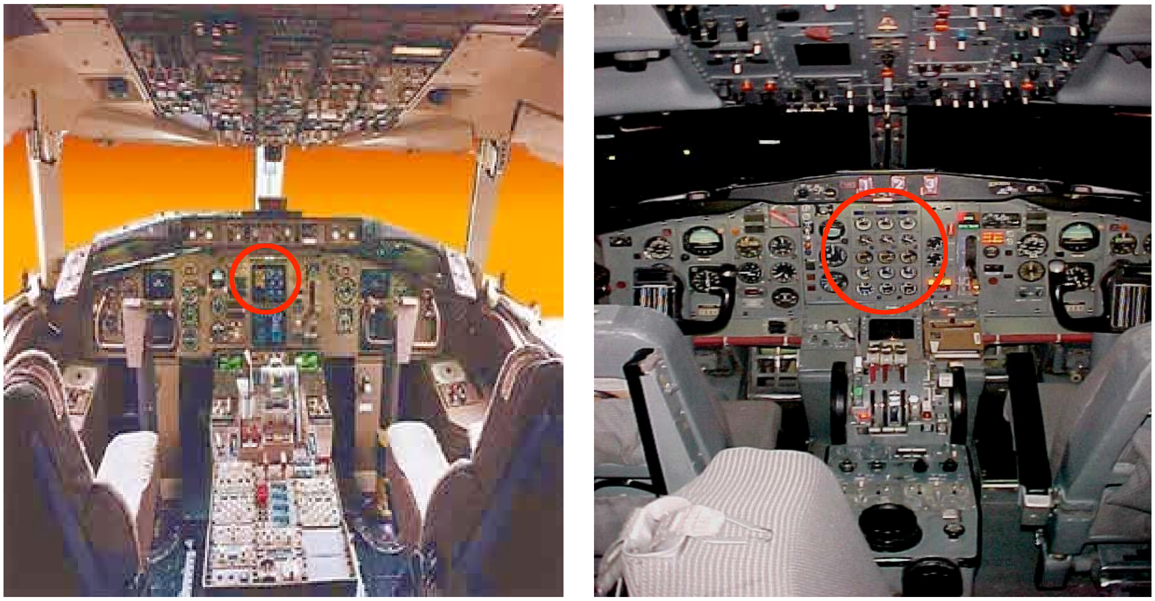


Figure 5. (a) B757 and (b) B727 flight decks (red circles compare engine instruments).

Because of a need to focus on specific displays and controls for approaches, we refer the reader to a web site to learn more about the B757 flight deck, if interested. The web site is: http://www.meriweather.com/767/767_main.html. Even though this site presents a B767 flight deck, the B757 and B767 are identical in layout (but the 767 is larger). The web site has a feature that allows a visitor to point to specific items in the photo to learn more about that item.

Figure 6 shows the layout of the most important flight deck displays. The reader should note the orientation of these important displays within the cockpit (Figure 5a). First we describe the controls and control modes, then the key displays. The key displays pointed to in Figure 6 are the primary flight display (PFD), navigation display (ND), airspeed indicator, altimeter, and engine instruments. Figure 6 also points to the gear handle (for lowering and raising the landing gear, or wheels) and flap position indicator – two other important items whose roles are further amplified, later, in the task analyses.

3.3.1.1. Flight Control Modes

There are three basic modes to control flight path: manual, automated, and partially automated. Manual flight is when the pilots steer the aircraft using the control wheel (or *yoke*) located in front of each pilot. The yoke is analogous to a car steering wheel, except that pilots also push or pull the yoke to descend or climb, respectively. The pilots also control speed via the engine throttles, located in between the pilots on the center console. To fly an approach manually, the pilots either look out the windows, if flying in good visibility, or use their instruments for guidance when steering their aircraft toward the desired landing spot on the runway.

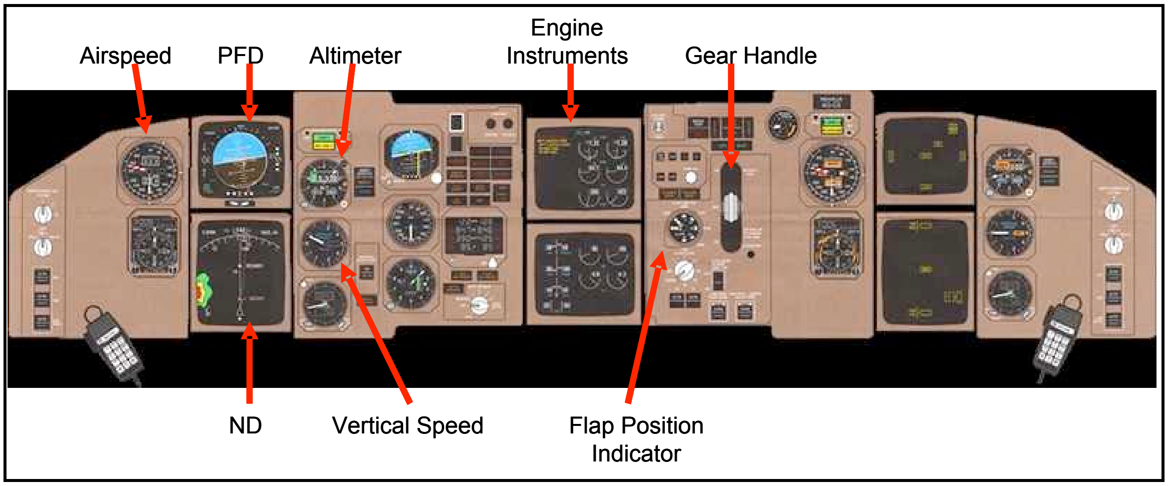


Figure 6. B757 center instrument panel, with labels for key items.

In automated flight, the pilots select the guidance method first (e.g., ILS or RNAV), and then engage the autopilot and auto-throttles to maintain the desired lateral and vertical paths, and speed. Even when in fully-automated flight control, the pilots have their hands resting on the controls in the event of a malfunction that would require their immediate response. Pilots also loosely hold the controls so that when they decide to land, they already have the feel of the control positions, as set by the autopilot, so they can more smoothly fly the landing.

In partially-automated flight, the pilots might use the autopilot, but not the auto-throttles, and so control airspeed themselves via the engine throttles. Or they might use only the auto-throttles for speed control and fly lateral and vertical paths themselves. Another option is to use the autopilot and auto-throttles, but steer using manual entries into the guidance functions via the mode control panel (MCP), described next.

3.3.1.2. Mode Control Panel (MCP)

The pilots use the MCP to select control and guidance modes for changing the aircraft path as needed. It is located at the top of the front instrument panel. As such, it has all the necessary controls for selecting flight control modes (automated to manual) and guidance cues in an up-front central location with easy access for both pilots (Figure 7).

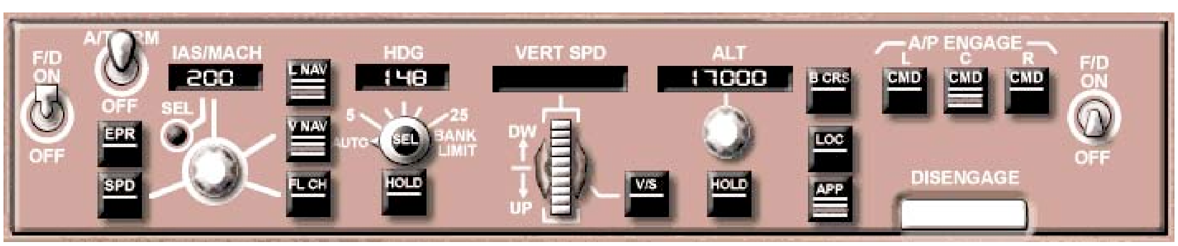


Figure 7. Mode Control Panel (MCP).

Each pilot has a flight director switch for turning on or off the flight director (F/D), which provides lateral and vertical steering guidance to the pilots on their primary flight display (more about the F/D in the PFD section, Section 3.3.1.3). Then, working from left to right, is the auto-throttle (A/T) switch for turning the auto-throttles on or off. Next is the airspeed or mach control. The pilots rotate the dial to select the desired speed. Buttons around the dial are for selecting the manner by which speed is controlled. EPR is engine pressure ratio, which gives units of thrust. Selecting EPR means that the pilots wish to control speed by thrust only. SPD is speed, which means that the pilots want to control speed by any method (engine thrust or pitch changes). The LNAV button selects the RNAV's lateral navigation for lateral guidance. The VNAV button is for choosing RNAV vertical path guidance. The bottom button in that column is FL CH, which is flight level change, a mode for climbing or descending to a specific altitude. In our approach examples, the pilots might use FL CH during the initial approach segment when descending from, say, 5000 to 3000 feet. The FL CH mode would use idle thrust and a moderate nose-down pitch to descend; airspeed could increase to any value in FL CH.

Moving to the right is the HDG, or heading controls. Pilots dial the desired heading (usually as instructed by ATC) using the middle knob where the bank limit is also set. Pushing the HOLD button holds the selected heading, overriding LNAV.

The middle thumb wheel, VERT SPD, allows the pilots to set a specific climb or descent vertical speed. For most approaches, the pilots could set a vertical speed of 500 to 1000 feet per minute, but are more likely to use VNAV or approach modes.

The next window, ALT, is for setting the desired altitude. As with heading (HDG), there is a button to HOLD the set altitude, thus overriding VNAV.

The next column of three buttons is for specific approach guidance. Ignoring B CRS, we will turn to LOC, which the pilots use for localizer guidance. The bottom button, APP, is for both localizer and glide slope (i.e., ILS), or LNAV and VNAV (i.e., RNAV) approaches. The three horizontal buttons are for selecting which autopilot (A/P) to engage. Redundancy is required for approaches where the A/P flies to landing. The bar beneath the A/P buttons is for quickly disengaging any engaged autopilot(s).

As a quick review, Figure 7 shows that LNAV, VNAV, Approach, and the center autopilot are selected. Plus, the auto-throttles and left flight director are on. Speed is set for 200 knots, heading is set to 148°, and altitude is set for 17000 feet. However, these settings are not guiding the aircraft because the pilots have selected LNAV, VNAV and Approach, along with the A/T and center A/P. So, the autopilot ignores the settings in the three windows and flies the LNAV and VNAV paths and speed. These settings could be used to prepare for the missed approach, which in this example would be to climb to 17000 feet at 200 knots on a heading of 148°.

Because the guidance modes that are engaged or armed on the MCP can be difficult to decipher based on a quick glance at the MCP, the armed and engaged modes are also annunciated on the primary flight display, described next.

3.3.1.3. Primary Flight Display (PFD)

The PFD is the primary attitude instrument. Both the captain (left seat) and first officer (right seat) have a PFD. The information provided by the PFD (Figure 8) is as follows.

Figure 8. Primary flight display (PFD).

Center of display:

- Artificial horizon depicted by blue and black ball. The blue half represents the sky; black is the ground.
- Transparent "aircraft wings" (outlined in white) depict the current attitude of aircraft in terms of pitch and roll.
- "Yellow cross" (looks pink-orange in the figure) depicts the Flight Director (F/D) command bars, which shows pitch and roll commands generated by the currently selected guidance source (e.g., ILS or RNAV). Typically, the autopilot or pilot rolls and pitches the aircraft to align the "aircraft wings" with the F/D command bars.

Upper left corner:

- GS200 – Ground speed in knots (200 knots in this example).

Upper right corner

- DH150 – Selected decision height or altitude in feet AGL (150 feet in this example).
- 1750 – Current AGL altitude in feet provided by radio altimeter (1750 feet in this example). The radio altimeter uses reflected radio waves from the ground to determine the height of the aircraft above the surface.

For the next two groupings, the depicted abbreviations are called flight mode annunciations (FMAs). FMAs in green font indicate the associated mode is *engaged* (active). FMAs in white font indicate the associated mode is *armed* and will engage under normally expected conditions when the *capture* conditions occur. For example, LOC armed means the localizer guidance will capture and engage when the aircraft enters the localizer cone (bottom half of Figure 3).

Lower left corner (for speed and vertical path modes):

- A/T (in green) – this location on the display indicates auto-throttle system status. In this example, the auto-throttles are engaged.
- SPD (in green) – this FMA means the pilots are using the auto-throttles to hold a selected speed (as opposed to flying a selected vertical path regardless of airspeed).
- G S (in white) – this location is for pitch control modes. In this example, G S (for glide slope) is armed, which means the glide slope will capture and engage when the aircraft flies within the glide slope cone (top half of Figure 3).
- V NAV (in green) – this FMA means VNAV is engaged and providing guidance for vertical navigation.

Lower right corner (for autopilot and lateral path modes):

- CMD (in green) – the autopilot or flight director status. In this example, CMD means the autopilot is actively flying the aircraft. If FD is displayed instead, it means the flight director command bars are displayed on the PFD and the autopilot is disengaged. If blank, the autopilot is disengaged and the flight director is off.
- LOC (in white) – this location on the display is for roll mode. In this example, LOC (for localizer) is armed.
- LNAV (in green) – this FMA means that LNAV is engaged and providing guidance for lateral navigation.

Bottom center:

- White dots and pink marker – localizer pointer (pink) and scale (white) indicate the center of the localizer beam with respect to the aircraft. In this example, the pink marker is right of center so the aircraft needs to turn to the right to fly to the center of the localizer cone. This is consistent with the F/D, which is commanding a turn to the right.

Center right:

- White dots and pink marker – glide slope pointer (pink) and scale (white dots) indicate glide slope position with respect to the aircraft. In this example, the pink indicator is above the center mark so the aircraft is below the glide slope. This is also consistent with the F/D, which is commanding a pitch up. A term often used by pilots is *one dot below glide slope*. This term refers to the pink indicator pointing at the first white dot *above* the center mark, but the aircraft is actually *below* the glide slope. On initial glide slope intercept, pilots use the "one dot below" indication as the latest time to configure the aircraft for the final approach if on the GSIA. Glide slope intercept usually occurs shortly thereafter.

Center left:

- White diamonds and pink marker – The fast/slow indicator depicts deviation from the pilot-selected or VNAV-directed airspeed. In this example, the indicator is centered, so no speed adjustment is needed.

Because the control modes and abbreviations can be confusing, Table 1 gives an overview of the MCP modes and corresponding FMAs. The MCP allows guidance modes to be either **engaged** or **armed**. A guidance mode that is engaged means that the guidance mode is actively providing signals to the displays, autopilot, or auto-throttles. A guidance mode that is armed means that the guidance mode will engage (i.e., become active) when the required conditions for its engagement have been met (e.g., intercept the localizer cone).

Table 1. Typical flight mode annunciations for approach (adapted from Casner, 2001).

Guidance Function	**How it works**	**FMA on PFD** (see Section 3.3.1.3)		
		Roll	**Pitch**	**Thrust**
HEADING SELECT	Roll used to maintain heading dialed into "HDG" window on MCP. 1) Dial new heading.	HDG SEL		
ALTITUDE HOLD*	Pitch used to maintain present altitude. 1) Dial desired altitude. 2) Push altitude "HOLD" button to maintain present altitude.		ALT	
SPEED	Adjustments to thrust used to maintain speed dialed into "IAS/MACH" window on MCP. 1) Dial new speed. 2) Push "SPD" button.			SPD
LOCALIZER*	Roll used to fly selected localizer course. 1) Dial ILS course and frequency on ILS panel. 2) Arm function by pushing "LOC" button. 3) Function captures localizer.	LOC		
APPROACH* (localizer + glide slope)	Roll used to fly localizer. Pitch used to fly glide slope. Adjustments to thrust used to maintain speed dialed into "IAS/MACH" window on MCP. 1) Dial ILS course and frequency on ILS panel. 2) Dial new speed (if needed). 3) Arm function by pushing "APP" button. 4) Function captures localizer and glide slope.	LOC	G S	SPD
FLIGHT LEVEL CHANGE	Thrust of engines set to idle. Pitch used to maintain speed dialed into "IAS/MACH" window on MCP. 1) Dial new altitude. 2) Dial new speed (if needed). 3) Push "FL CH" button. 4) Descends to new altitude and then switches to ALTITUDE HOLD.		FLCH SPD	HOLD

* Guidance functions that can be armed prior to engagement.

As can be seen in the Table 1, the HEADING SELECT, ALTITUDE HOLD, and SPEED guidance functions are each dependent on only a single state – roll, pitch, and thrust, respectively. The commands for HEADING SELECT and SPEED come directly from pilot entry into the MCP. For example, if HEADING SELECT is engaged, the aircraft will begin to turn as soon as the pilot changes the heading value in the "HDG" window on the MCP (Figure 7).

The command for ALTITUDE HOLD comes from two possible sources. In the first case, by pressing the "HOLD" button under the "ALT" window on the MCP, the aircraft will hold the current altitude indefinitely. In the second case, the altitude entered in the "ALT" window on the MCP becomes the target altitude. However, in this latter case, in order for the aircraft to change to the target altitude, FLIGHT LEVEL CHANGE must first be selected. When FLIGHT LEVEL CHANGE is engaged, the ALTITUDE HOLD function is said to be *armed*. In this case, when the aircraft descends and reaches the target altitude, the armed condition is met and the guidance function disengages FLIGHT LEVEL

CHANGE and engages ALTITUDE HOLD. In addition, the engaged pitch FMA on the PFD switches from "FLCH SPD" to "ALT" (not shown in Figure 8), and the engaged thrust FMA on the PFD switches from "HOLD" to "SPD." (The terms *flight level* and *altitude* are used interchangeably in this context).

Another guidance function that is armed prior to being engaged is APPROACH. As an example, consider an aircraft flying with constant heading, altitude, and speed via HEADING SELECT, ALTITUDE HOLD, and SPEED functions. If APPROACH is armed, it becomes engaged when the aircraft intercepts the localizer (assuming, of course, that it is on an intercept course). The engaged roll FMA switches from "HDG SEL" to "LOC." At this point, if the autopilot is engaged, the aircraft flies the course corresponding to the localizer. When the aircraft intercepts the glide slope, the engaged pitch FMA switches from "ALT" to "G S," corresponding to aircraft commands to fly the glide slope. The thrust FMA remains unchanged displaying "SPD."

3.3.1.4. Navigation Display (ND)

The ND provides the pilots with a map view (Figure 9) of the area in which the aircraft is headed. Both the captain and first officer (FO) have an ND. The ND can be configured in various modes with *map* mode being the most common. In fact, during approach, it is common for the ND of one of the pilots to be in map mode and the other pilot to be in ILS mode. ILS mode allows the raw ILS signal data to be displayed. Using different modes allows the pilots to crosscheck information. For example, the map mode displays information based on the RNAV-computed position. If the aircraft location is in error for whatever reason (e.g., drift in the inertial navigation system), there would be no way to know this from the map mode. However, if the ILS mode is being used by the other pilot and there is an ILS signal detected, then the discrepancy would become apparent by crosschecking the two displays. Details of the ND, by region of the display, are as follows.

Bottom center:

- White triangle – aircraft symbol. The apex of the triangle indicates the aircraft position relative to display.
- White dashed line – curve trend vector. Indicates predicted airplane track in 30, 60 and 90 second intervals when turning.

Center:

- "AMBOY" and "KTTN" with waypoint symbols – indicate waypoints. White for inactive, magenta for active (there is no active waypoint shown in the figure).
- "SOJ" with VORTAC symbol – indicates VORTAC navigation aid. When the NAVAID switch is on, all appropriate navigation aids within the range scale appear, in addition to those navigation aids which are standard or active to denote fixes.
- "MMP" with blue circle – indicates the MMP airport. (Airports have unique three-letter identifiers; e.g., LAX. In this example MMP is fictitious.) When the ARPT switch is on, the ND displays airports within the map area.
- Pink solid line with tic marks – indicates the course based on a prediction using the present heading and wind.
- "80" – the range from the aircraft to the associated tic (80 nm in this example). Also indicates half of the range selected on the ND range selector. In this example, the ND range selector is set to 160 nm, displaying a moving map 160 nm in front of the aircraft.
- Magenta solid line – indicates planned lateral route of flight.
- Pink dashed line – indicates the heading set in the MCP. In this example, the heading is set to 035 degrees.

Figure 9. Navigation Display (ND).

- Green arc – altitude arc. The intersection of the green arc with either the course line or flight plan route predicts the point where the aircraft will be at the MCP altitude, assuming no changes to the current attitude and vertical speed.

Upper left corner:

- 4.4 NM – indicates the distance to the active waypoint (4.4 nm in this example). In the figure, the active waypoint is not shown and may be hidden beneath the aircraft symbol.

Upper center:

- TRK 062 M – magnetic track display (062 degrees in this example). The true heading is displayed by a white triangular pointer along the compass arc. In this example, the true heading is 073 degrees. The difference between true and magnetic is due to the difference in location between the Earth's true North Pole and its magnetic north pole. The angular difference between the two poles, which varies based upon the aircraft location, and so is displayed on the ND.

Upper right corner:

- 0835.4z – indicates the estimated time of arrival to the next waypoint in Zulu (Greenwich Mean Time).

Bottom right corner:

- Vertical white line and tics with pink diamond – the vertical deviation indicator, which depicts the altitude deviation from the RNAV's vertical profile (VNAV path). In this example, the airplane is slightly below the desired VNAV path.

Bottom left corner:

- White arrow and "120" – the wind display. Indicates wind direction relative to map display orientation, and speed in knots. In this example, the aircraft is experiencing a 120-knot right-quartering tailwind.
- Multicolored blob – weather radar symbol. The highest intensity is displayed in red, lesser intensity in amber, and least intense in green. Turbulence is displayed in magenta (not shown in this example).

Center left corner:

- TFC – means the TFC button is on, which means the TCAS traffic display is on.

3.3.1.5. Miscellaneous Controls & Displays

There are a few other systems that are relevant to approaches and our objectives: the flight management system, the traffic collision avoidance system, and two types of altimeters.

<u>Flight Management System</u>

The flight management system (FMS) is one instance of an RNAV system. The FMS assists the pilots in planning and executing the flight route. During flight planning, the pilots enter the flight route, aircraft weight, and expected weather conditions into the FMS via the central pedestal control display unit (Casner, 2001). Information about the flight route includes the *expected* departure runway and departure procedure, cruise speed and altitude, arrival and approach procedures, and the *expected* runway assignment. The *actual* flight route can differ from the plan (and oftentimes does) depending on weather and ATC requirements. Changes enroute require the pilots to re-program the FMS. The FMS calculates the optimal flight path and economical speeds during the climb, cruise, and descent phases of flight. The FMS-calculated lateral path is presented on the ND as the magenta line between waypoints (in Figure 9).

Although the FMS paths can theoretically be followed from takeoff to just prior to landing, the reality is that ATC clearances during the descent and approach phase of flight often differ from what has been programmed into the FMS. Pilots do not typically re-program the FMS to account for ATC clearances just prior to or during the approach for two reasons. First, reprogramming the FMS requires a significant amount of time and cognitive workload, and it distracts the pilot doing the programming from other important tasks, such as monitoring the aircraft position, speed and configuration (Degani et al, 1995). Second, ATC clearances just prior to or during the approach do not typically require complex lateral or vertical path planning. Rather, these ATC clearances instruct the aircraft to change heading, altitude, and speed (or any combination of the three). Therefore, pilots use the MCP to comply with pre-approach clearances (i.e., during initial or intermediate segments).

<u>Traffic Alert and Collision Avoidance System</u>

The traffic alert and collision avoidance system (TCAS) is an airborne collision avoidance system based on each aircraft's ATC radar beacon signals. TCAS operates independent of ground-based equipment, and gives the pilots traffic advisories and collision avoidance guidance. TCAS is a backup system to ATC, which is the primary system for keeping aircraft safely separated.

Altimeters

Commercial jet aircraft have two types of altimeters, barometric and radio. Barometric altimeters function based on relative air pressure. They are generally pressure corrected to measure the altitude above mean sea level (MSL). Radio altimeters, on the other hand, measure the absolute distance to the ground by transmitting a signal from the aircraft to the ground and measuring the time to receive the reflected signal. While the pressure-based altimeter is used during all phases of flight, the radio altimeter is primarily used during the approach and landing phases to measure the exact distance above ground level (AGL). Most radio altimeters are active when the aircraft is below 2500 feet AGL; it is inactive above this altitude.

Since we have explained the relevant controls and displays for ILS and RNAV approaches during the above discussion, we now turn our attention to the unique synthetic vision display concept.

3.3.2. Synthetic Vision System (SVS) Display

This section provides a detailed description of SVS as it applies to approach and landing, and a discussion of the concept of operations for the system. It is important to realize that, as a developing product, the details of system implementation will change as experiments and tests are completed and new system components are integrated. As such, the system components and detailed descriptions provided here represent only a snapshot in time of system configuration and implementation. In addition, multiple companies are developing SVS concepts. This document uses examples from two different concept designs.

The SVS display includes terrain and airport details and may include other aircraft, weather and wake turbulence information. In addition, flight data information is overlaid on the terrain background along with a tunnel navigation aid that presents a "highway-in-the-sky" for the pilots to follow. Figure 10 shows one example SVS terrain display with a flight data overlay (without the tunnel).

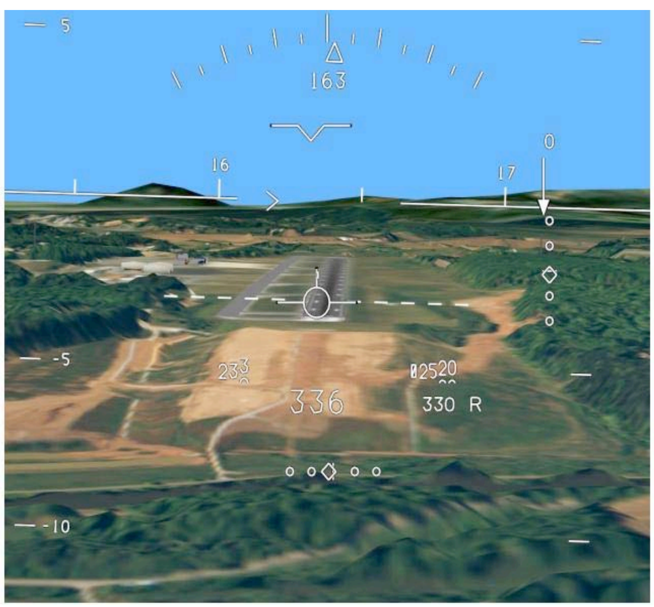


Figure 10. SVS terrain and airport display with flight data overlay.

Following are descriptions of the concepts and display components for the system. The purpose here is simply to present some of the differences in the design concepts to give the reader an idea of the range of possibilities being explored, and to illustrate the similarities. It should be understood that the concept designs will change as the NASA AvSP program continues.

Terrain Map

The terrain map is the computer-generated image of the terrain from the pilot's viewpoint. It is expected that the image will be highly intuitive as it replicates what a pilot would see out the front window in good visibility daylight conditions. The image is generated by the SVS sensors and a terrain database, and can be implemented in a number of different ways. In one of the NASA concepts, a technique called photo-realism is used to combine the terrain database information with high-resolution photos of the area (Figure 11a). The Rockwell Collins concept uses a graphical texturing technique and terrain altitude information to create the image (Figure 11b).

Flight Data

The flight data overlays include speed and altitude information on vertically oriented bars on either side of the display, and heading on a compass arc at the top of the display. Approach guidance information for horizontal and vertical path orientation and an artificial horizon are also included. These PFD-like images do not differ much between the two example SVS concepts.

Velocity Vector

The velocity vector (pointed to by red arrows in Figure 11), also called the flight path symbol, represents the actual flight path of the aircraft. Flight control inputs by the pilot or autopilot cause the symbol to move, providing instant feedback. In essence, the pilots use the aircraft controls to fly the velocity vector in the display to the desired location. While the velocity vector is not available on current B757 instruments, it is available on other aircraft. A velocity vector is implemented in both concepts discussed here.

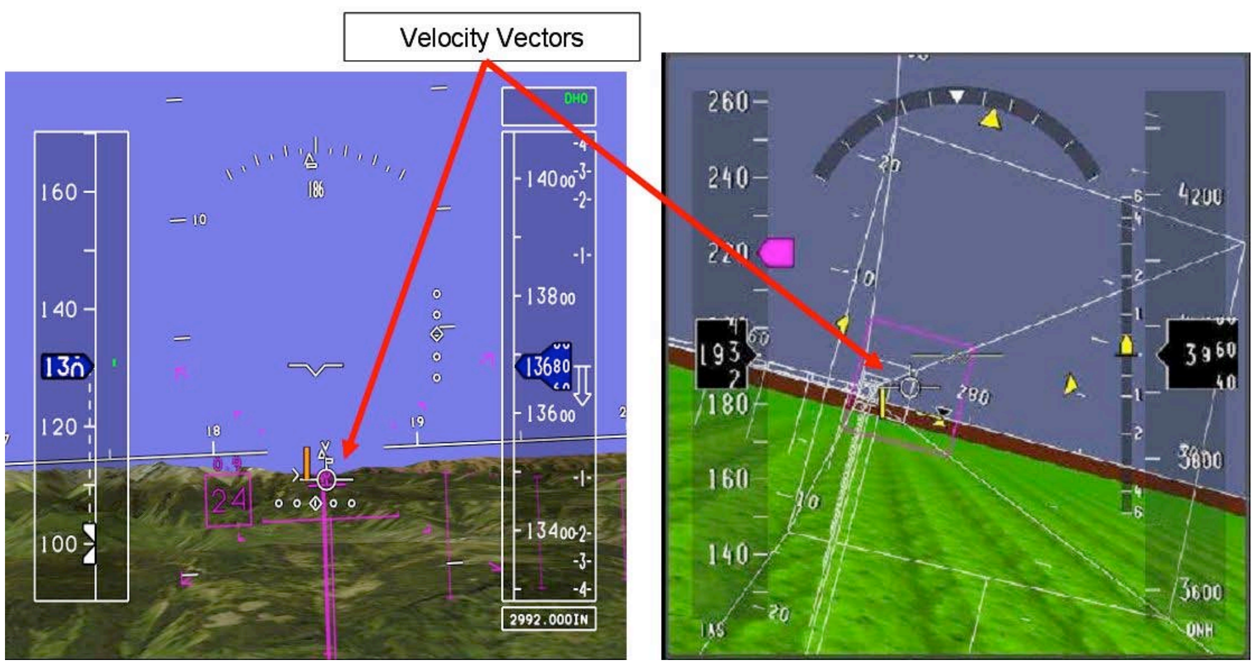


Figure 11. (a) NASA SVS concept and (b) Rockwell Collins concept.

Tunnel Navigation

The tunnel navigation or highway-in-the-sky is a flight path guidance concept. Graphics on the terrain display show the desired flight path as it extends out in front of the aircraft. The NASA tunnel graphic uses small magenta brackets to indicate the corners of the tunnel and a T shape or *goal post* to indicate the bottom of the tunnel and distance to the ground. It also uses a graphic of an aircraft, sometimes called a *ghost* aircraft that travels ahead of the aircraft along the path of the tunnel. These tunnel and ghost aircraft symbols are very subtle and difficult to see in Figure 11. The Rockwell Collins tunnel graphic uses connected squares that reduce in size to present the image that the tunnel extends ahead of the aircraft. A magenta colored box moves ahead of the aircraft along the path of the tunnel. The Rockwell Collins concept does not have a ghost aircraft. It only requires that the velocity vector be inside the magenta box. The NASA concept uses the ghost, but it is difficult to see in this image as the velocity vector is right on top of it.

3.3.2.1. Concept of Operations

The designers of SVS have defined a concept of operations for the system that ranges across all phases of flight and focuses on safety and improved operating efficiency (Williams et al., 2001). This section presents a basic overview of this concept followed by a more detailed concept of operations for the approach phase of flight.

The intent of the SVS is that it allows aircraft to operate in low visibility conditions with the same or greater level of safety and by similar rules as those used during conditions of good visibility. Visual Flight Rules (VFR) assume that a pilot can see other nearby aircraft. As such, the rules allow for reduced aircraft spacing. This translates into a greater operational tempo as more aircraft can occupy a given volume of airspace. Also, the pilots can see the airport and any obstacles to the approach path at a much greater distance. In low visibility conditions the flight rules change to Instrument Flight Rules (IFR) to reduce the risk created by the loss of visibility. IFR requires increased aircraft spacing and specific airport-aircraft instrument combinations to continue operations. As such the operational tempo is reduced (Williams et al., 2001).

The SVS concept of operations supports the operational tempo and safety levels of VFR during IFR conditions down to the lowest visual minimums, while using non-ILS equipped airports and runways. In other words, airports that are not currently equipped with ILS systems should be able to use VFR aircraft spacing in very low visibility conditions, if the aircraft are equipped with SVS.

In addition, the SVS concept includes path guidance and hazard avoidance. Path guidance (e.g., the tunnels described above) enables the pilots to follow a desired flight path. The tunnel navigation concept of current SVS designs is focused on improving path control beyond current instrument guidance systems. The focus on hazard avoidance is one of the critical safety elements of SVS. The system should increase the ability of pilots to detect, identify, prioritize, and avoid hazards, which includes traffic, terrain, obstacles (e.g., radio towers), wildlife (e.g., a flock of birds), and weather.

In addition to the potential change to aircraft spacing rules, SVS supports several other approach operations (Williams et al., 2001). At airports where ILS or other approach system components are inoperative or unavailable, SVS could support approaches using VFR criteria. An example might be the ability to use lower visibility minima when approach lighting or the glide slope signal is inoperative. SVS could also be used to augment current instrument systems either as an independent check of accuracy, or as an addition that supports the use of lower visibility minima than would otherwise be allowed based on a published ILS or RNAV approach for a given runway.

Approach path control will be enhanced by the tunnel navigation component of SVS to include *circling,* and published visual approaches. Circling approaches involve flying an approach to one runway, then aligning with and landing on a different runway. Published visual approaches follow specific terrain features, such as flying over a river to keep aircraft noise away from populated areas. Both of these approach types, which are normally only flown in good visibility, could be performed when visibility is poor, using an SVS.

SVS would also help mitigate the serious problem of controlled flight into terrain. It is presumed that, given a view of terrain hazards similar to daytime VFR, pilots will avoid flying into a hillside, mountain, or other terrain, as sometimes tragically occurs during conditions of reduced visibility (e.g., due to fog or heavy rain). Pilots will also use an SVS (and its velocity vector) to avoid landing short of, or beyond, the desired touchdown point on the runway.

The location of the SVS display is open for debate and depends upon retrofit capabilities of various aircraft. One prototype has the display in front of the left seat pilot only (Figure 12).

3.3.3. Displays & Controls Summary

The cockpit controls and displays, when combined with radio information and out-the-window observations, provide all of the needed information for pilots to safely plan and fly the desired flight path. There is one last detail of background information to mention before presenting the pilot tasks that serve as the foundation of the human performance models. This last detail is flight crew responsibilities.

Figure 12. SVS in a NASA test aircraft.

3.4. Flight Crew Responsibilities (PF and PNF)

The B757, like most modern commercial jet aircraft, has a two-pilot flight crew. The crew consists of a captain and first officer. The captain sits in the left seat, and the first officer sits in the right seat. The aircraft can be flown from either position. The pilot currently flying the aircraft is known as the **pilot flying** (PF). The other pilot is the **pilot not flying** (PNF). Pilots exchange duties throughout a typical day of flying to multiple destinations so that both pilots maintain their skill levels through all phases of flight. Air carriers establish crew responsibilities and coordination procedures to keep both pilots actively aware of the aircraft position and condition throughout the whole flight.

Normally, the PF actually manipulates the flight controls or the autopilot control modes (via the MCP), while the PNF monitors progress, reads checklists, moves gear and flap levers (as instructed by the PF), and communicates with ATC and the flight attendants. The idea is to minimize the distractions for the PF, but keep the PNF sufficiently aware of the aircraft state so that, in the event of an emergency, either pilot can assume aircraft control. If the PF is hand-flying the aircraft (i.e., not using the autopilot), then the PNF will also set values and modes on the MCP. The reason is that hand-flying takes two hands: one on the control yoke, the other on the throttles. Such a situation occurs during landing when the PF is busiest ensuring that the airplane touches down on the desired spot on the runway. Occasionally, the PF opts to use the auto-throttles while hand-flying, in which case one hand might be free to set MCP modes, for example. Even then, the PF's "throttle hand" (right hand for the captain; left hand for the first officer) stays near the throttles in case of a malfunction.

4. Approach Task Analyses

Each approach type (visual, ILS, RNAV and SVS) has similar tasking for the pilots in terms of human performance. A common taxonomy for analyzing human performance is the VACP method (Wickens, 1984). The VACP method examines an operator's **V**isual, **A**uditory, **C**ognitive and **P**sychomotor (manual) workload. Briefly, the visual component for pilots is to visually monitor their instruments and acquire the runway. The auditory component is for communications between the pilots, to and from the cabin crew, to and from ATC, as well as any cockpit system annunciations. Pilot cognitive workload is due to interpreting the observations of their instruments and communications, deciding how much automation to use, and deciding to land or execute a missed approach. The last component, psychomotor, refers to manipulating automation or flight controls to achieve the desired approach path. Also, actual landings are manually flown.

This section discusses the tasks and key decisions for the approach phase of flight for the B757. The specific information, presented in sub-sections, covers the following:

- 4.1 General and ILS approach tasks
- 4.2 RNAV approach tasks
- 4.3 ILS and RNAV task timelines
- 4.4 Cognitive decisions
- 4.5 Typical approach problems and errors
- 4.6 Information requirements and situation awareness
- 4.7 SVS approach benefits

4.1. General and ILS Approach Tasks

The task analyses focus on the approach phase of flight. Following the cruise and descent phases (Figure 1), the pilots transition to the approach for a specific airport. The approach is the portion of the flight during which the pilots fly the aircraft into the appropriate location, attitude, and configuration to land. For both manual and automated flight, this involves incrementally slowing to landing speeds, descending to appropriate altitudes for landing, and aligning the aircraft with the runway such that the landing can be executed at the correct attitude (wings level and within pitch tolerances) and correct speed (within tolerance), within the appropriate runway touchdown zone. The maneuvers performed by the crew for both the approach and landing must be within the limitations of the aircraft, the procedures of the airline, and the requirements of ATC, while ensuring the safety and comfort of any passengers.

During the approach, the pilots make a series of speed reductions and wing flap deployments in order to maintain the necessary pitch window and descent rate of about 300 feet per mile (3 degrees) to land at the correct speed. Landing at too fast a speed requires too much wheel brake and reverse thrust energy to stop the aircraft; landing at too slow a speed risks stalling the aircraft and crashing. Slowing to landing speed requires the use of flaps to maintain pitch tolerances. Without flaps, the pitch has to be too high (nose up) to be safe at the slower landing speeds, and again risks a stall. Also, as flaps are lowered, the pilots maintain certain speed ranges to avoid over-speeding the flaps or flying too slow for that increment of flap setting. A minimum flap setting is a function of the weight and airspeed of the airplane, so that, as the airplane slows toward landing speed at a given weight, progressively greater flap settings are required. Flap settings for a typical B757 landing weight of 180000 pounds are in Table 2. (This data was provided by one of the subject matter experts, a United Airlines B757 pilot.)

Table 2. Typical B757 flap settings.

Airspeed	**Minimum flap setting**	**Maximum flap setting**
(knots)	f(weight, speed)	f(speed only)
240	0	1
220	0	5
210	0	15
205	1	15
195	1	20
185	5	25
165	15	25
145	20	30
125	30	30

During the initial approach segment, the crew also sets the guidance and control systems based on the type of approach for the selected runway. In the case of an ILS approach, the settings enable the autopilot to intercept the runway localizer and glide slope guidance signals. In addition, the crew communicates with ATC as needed to obtain appropriate altitude, approach and landing clearances, and other information and instructions. The crew also monitors the aircraft sub-systems, position, attitude, and flight path.

The tasks are shared (usually by established procedure) between the PF and the PNF. The primary tasks of the PF involve all aspects of the aircraft attitude, position and path control. The PNF performs necessary communications, sets required navigation and radio frequencies, accomplishes checklists, responds to requests from the PF for flap and gear deployments, and double-checks PF actions. Usually, tasks of monitoring the displays and looking out the windows for traffic, obstacles, and the runway are performed by both pilots.

For the purposes of the ILS approach task analysis, we made several basic assumptions about the aircraft and set several initial conditions for the beginning of the approach phase tasks. The following are the assumptions and initial conditions of the aircraft and flight, and a brief narrative of the approach and landing task sequence.

For this scenario, it is assumed the B757 aircraft weighs 180,000 pounds. The associated speed and flap settings are listed with the task table. The aircraft has finished the descent phase of flight and has begun the approach phase. The flight is proceeding under IFR towards an ILS runway. The aircraft is approximately 15 miles from the runway threshold flying at 3000 feet AGL at 200 knots, with flaps set at 5 degrees. The pilots have intercepted the localizer and the tasks begin as they receive approach clearance.

Upon receiving the approach clearance from approach control and instructions to slow to 180 knots, the PF sets the MCP approach mode, the MCP speed to 180, and calls for flaps 15. Once the PF sees the glide slope pointer begin to move from the top of the scale (known as "glide slope alive"), the PF calls for the landing gear, reduces speed again, and calls for flaps 20. Before glide slope capture (usually corresponding to what pilots refer to as "one dot below glide slope") the PF again reduces speed and calls for flaps 25. As the aircraft intercepts the glide slope, the PF reduce speed again and directs the PNF to set flaps 30 (the final flap setting), which completes the Before Landing Checklist.

After glide slope capture, the crew sets the missed approach altitude and heading. Upon crossing the outer marker at about five miles from the runway threshold, the pilots confirm their final approach preparations, and the PNF changes the radio to the tower frequency. Upon changing frequency, the PNF calls the tower and receives landing clearance. Approaching decision height, the crew makes their landing decision. If they cannot see the runway, they perform a missed approach. If they see the runway and decide to land, the PF announces that decision aloud and normally takes manual control of the aircraft. If the runway is in sight before the decision height is reached, the PF will often switch to manual flying at that sighting. The point at which the PF begins to manually fly the aircraft varies, but is usually associated with the ability to see the runway.

Detailed task descriptions follow for all approaches and for the ILS approach specifically. After these task descriptions, we describe the unique RNAV approach tasks, and then we present the baseline (ILS and RNAV) tasks in a timeline table.

4.1.1. General Approach and ILS Task Descriptions

First we describe the sequential tasks, then the non-sequential tasks. The distinction between sequential and non-sequential tasks is important to human performance modeling and will affect how the specific tasks are represented within the models

4.1.1.1. Sequential Tasks

<u>Communicate with ATC</u>

When the crew communicates with ATC, it is either initiated by the crew or in response to communication from ATC. Contact initiated by the crew usually takes the form of an identification call or a request for clearance or information. Responses usually involve reading back ATC instructions or providing requested information (such as present speed). Voice communication requires the PNF to press one of the microphone (mic) buttons on the yoke or on the center console while speaking into the PNF's headset microphone. To end a radio call, the PNF releases the mic button.

<u>Set Radio Frequencies</u>

The two radio control panels on the center pedestal between the two pilots allow for two communication radio frequencies to be selected. A toggle switch on the panel allows the crew to switch from one frequency to another. During the approach, the crew uses the approach control frequency. At or near the FAF, the PNF flips the switch to select the tower frequency.

<u>Engage Automated Flight Control</u>

Arming the approach mode, and engaging the auto-throttles and an autopilot, is how the PF selects automated flight control. These selections enable the autopilot to fly the localizer and glide slope until the pilots take manual control. If the PF does not choose automated flight, then the PF must follow the ILS guidance pointers or flight director to maintain the correct lateral and vertical paths. Arming the approach mode requires pressing the approach mode button labeled APP on the MCP; selecting an autopilot requires pressing the selected button, also on the MCP (Figure 7).

Maintain Airspeed

Maintaining airspeed requires looking at one of the two airspeed indicators. The indicators include movable markers along the outside of the dial called *bugs* that are set to reference speeds during approach preparations. The pilots set the bugs using checklist-like charts that list the target airspeeds for given aircraft weights and flap settings. The bugs are memory aids, since the correct speeds change from flight to flight with aircraft weight. The PF refers to the bugs to set the MCP speed during automated flight; the auto-throttles maintain the set speed. Setting the speed requires turning the speed dial until the desired speed is indicated by the digital display above the dial (Figure 7). Verifying that the correct speed has been set requires looking at this display.

Set Flaps

The flaps are usually set by the PNF but the flap lever is easier to reach from the right seat, since it is to the right of the throttle controls. When the flaps are set from the left seat, it requires more care to reach around the throttles, and the flap lever position is more difficult to determine. Setting the flaps requires using one hand to move the flap lever to the correct position and requires the pilots to look at the flap indicator on the front instrument panel (Figure 6) to confirm the correct flap position. The flap lever slides into a detent for each available flap increment (1, 5, 15, 20, 25, and 30 degrees). During approach, the PNF (usually) moves the lever to the next appropriate position when called for by the PF. Pilots may also feel the position of the flap lever with their throttle hand to check if the lever is settled into the correct detent.

Monitor Localizer and Glide Slope

As the approach continues under automated flight, both pilots monitor their PFDs to ensure proper following of the localizer and glide slope signals (Figure 8). They also monitor the FMAs and MCP to ensure that the correct modes are engaging after being armed.

Lower Landing Gear

Lowering the landing gear requires moving the landing gear lever all the way down. As with the flaps, the lever is closer to the right seat (Figure 6) because a default assumption is that the captain (left seat) is the PF, while the first officer (right seat) is the PNF. The gear lever requires only one hand to pull the lever out slightly and then push it down. If done from the left seat, it requires leaning the upper body toward the lever to reach it. Verifying that the gear is locked down requires looking at the three indicator lights directly above the landing gear lever. The lights are positioned in a triangle (nose, left and right gear). If all three lights are green, then the landing gear is, in pilot terms, *down and locked*. Both pilots check the gear lights, usually after they hear the gear lower into position with a distinct “thunk” sound, to be sure the gear are down and locked.

Arm Speed Brakes

The speed brakes are controlled using a lever on the left side of the throttles. The lever is moved back (aft) to deploy the speed brakes (or spoilers), which are panels on the top of the wings that spoil the lift of the wing and allow the aircraft to descend faster or slow down more quickly. The speed brakes also work automatically upon touchdown of all landing gear to slow the aircraft. In the forward position, the lever is in a detent indicating that the speed brakes are stowed (i.e., flush with the wing surface). The next aft setting is the *armed* position used for automatic deployment during the landing roll. Beyond that, the lever can be moved farther aft to vary the amount the spoiler panels are raised. Pilots use varying spoiler positions depending upon how quickly they wish to decelerate. To verify that the speed brakes are armed for landing during the final approach segment requires the pilot to look at the lever position, and to sometimes use one hand (again, the throttle hand) to feel that the lever is in the armed detent. The speed brake lever is easier to reach from the left seat, since it is left of the throttles. Deploying the speed brakes from the right seat, if the FO is the pilot flying, requires the FO to reach around the throttles.

Set Missed Approach Altitude

The missed approach altitude is the altitude to climb to in the event of a missed approach and is given on the approach chart. Setting this missed approach altitude requires using the altitude knob on the MCP to dial the correct altitude (Figure 7). Typically, the PF sets this altitude, which must be done after glide slope intercept, and the PNF verifies it.

Monitor Altitude below 2500 Feet AGL

The reading for the radio altimeter is on the PFD just below the decision height (DH) reading (Figure 8). The PNF calls out AGL altitudes, typically at 1000 and 500 feet, to denote the standard stabilization gates. Both pilots check for the aircraft being within stable approach parameters. The PNF also calls out “Approaching decision height” (usually 100 feet above) and “Decision height” if the PF has not yet indicated that the runway is in sight.

Before Landing Checklist

The Before Landing Checklist is a sequence of steps that are executed by the PNF which are designed to verify that certain critical tasks have been completed prior to landing. Each step is called out by the PNF and an associated check is done. Depending on the airline, the PF is not required to verbally respond to any of the checks as they are the duty of the PNF. However, the PF will usually follow along with the checks and verify each one as the PNF reads through the list. Most airlines require the landing checks to be done by reading the steps from the checklist card, rather than from memory, to avoid missing any critical item.

Turn on Landing Lights

The controls for the landing lights are three switches (for left wing, right wing and nose gear lights) on the middle overhead panel. The PNF turns the landing lights on while accomplishing the Before Landing Checklist, if not sooner. The lights are required to be on, unless using them is distracting (when, for example, it is night in the clouds and the reflected light would harm the pilots’ night vision).

Monitor Descent Rate

The descent rate is determined by looking at one of the two vertical speed indicators, which are analog dials showing the vertical speed in feet per minute (Figure 6). For a typical precision approach and B757 ground speeds, the vertical speed should be about 700 feet per minute to fly the desired glide path.

Disengage Autopilot

The PF disengages the autopilot via the MCP or by using a button on the control yoke when he or she decides to fly the aircraft manually. Prior to disengaging the autopilot, the PF will put both feet on the rudder pedals and place his or her hands on the yoke and throttles. Once the PF is ready to assume control, he or she will press the yoke button with his or her thumb. Or the PF may ask the PNF to press the MCP disengage bar (Figure 7), or do so him- or her-self. A cockpit alarm sounds as the autopilot disengages. If the PF uses the yoke button to disengage the autopilot (which is the typical method), then the PF presses that button again to silence the alarm.[3]

Fly Manually

Manual flight by the PF requires both hands, both feet, and visual scans of the instruments and out the front window. It also requires some attention to the radio and PNF who may call out information that the PF needs to know. Scan patterns vary from pilot to pilot, but most pilots spend roughly equal time looking out the window, and looking at the PFD and surrounding instruments during final approach. If visibility is poor, the PFD is the primary focus. The PF makes constant minor adjustments to maintain runway alignment, wings level, on speed, and the desired descent rate using the yoke, rudder pedals and throttles.

Flare

To flare the aircraft, the PF gradually pulls on the yoke when over the runway to bring the pitch up to the landing attitude, while reducing the thrust to idle on both engines. The flare also requires the PF to keep the wings level and the airplane aligned with the runway centerline while permitting the airspeed to decrease to touchdown speed (usually about 5-10 knots below final approach speed). An ideal flare to touchdown occurs when the pitch reaches the desired angle and the engines reach idle thrust as the main landing gear simultaneously contact the runway.

Next are the non-sequential tasks.

4.1.1.2. Non-sequential Tasks

Monitor Flight Path and Progress

This task is periodically performed by both crewmembers throughout all phases of flight. The task primarily involves scanning the instruments to ensure that the aircraft has not deviated from the expected path, altitude, attitude, airspeed and overall flight plan. Looking at the ND (Figure 9) allows the pilots to determine if the aircraft is on the desired flight path, as programmed into the FMS. Looking at the PFD and its FMAs allows the crew to verify that the aircraft is in the prescribed attitude and that the automated flight systems are functioning normally. Pilots mainly look out the windows, if visibility is good, to verify the correct airplane attitude. Other displays such as the vertical speed indicator allow the crew to monitor the progress of various changes or determine that unexpected changes may be occurring.

Double-Checks and Verifications

Throughout the approach and landing process both pilots check and double-check the accuracy of settings that include altitude, speed, and flaps. Sometimes these checks require consulting a reference such as the speed versus flaps settings based on the weight of the aircraft. Other times the same steps are done so frequently that the crew has expectations of what the settings will be. In these cases double-checks are more of a mental process of determining if an expectation has been violated. For example, if the PNF is expecting a particular flap setting and the PF asks for a different one, the PNF would query the PF to determine the reason for the difference.

Monitor the Radio

This task involves listening for communications on the current radio frequency. Auditory information is received through the ear piece, headphones, or cockpit speaker. The information may include specific communications from ATC directed at the crew, or communications between ATC

[3] Small, R.L. (5/23/03). Personal communication with Delta Airline's Captain Bill Jones, a B757 pilot.

and other aircraft. This monitoring task requires no workload when there is no communication traffic on the frequency because there is no information available to monitor. Attention is directed to the radio when the pilots initiate a transmission or when attention is drawn by communications on the radio. When communications do occur, the crew quickly determines if the information is directed at them based on their call sign (e.g., NASA-113 in the earlier examples). They also quickly determine if the communication is coming from ATC or from another aircraft. When the radio call is for the crew, they will closely attend to the information – even writing down clearances to ensure accuracy. The pilots also monitor communications between ATC and other aircraft because it helps them anticipate what ATC may direct them to do and how ATC is managing the airspace, especially during the approach phase. ATC calls to them will either confirm their expectations regarding approach and landing clearances, or require them to make some sort of change. Listening to communications from ATC to other aircraft helps the crew build a mental picture of where they are in the airspace relative to the other aircraft and provides them with an idea of what to expect as they get closer to the airport. The pace of radio communications will vary depending on a variety of factors including the weather and the quantity of aircraft approaching the airport. At its worst, the calls on the radio can be continuous as ATC and flight crews initiate calls and respond to each other, which require some level of constant attention by the pilots. At such times, it can be difficult to find a break in the communication flow to initiate a call. It is not unusual during such situations for multiple aircraft to "talk over" each other at the same time, which adds to the confusion and hectic tempo.

Monitor Aircraft Systems

This task is periodically performed by both pilots throughout all phases of flight. The status of all of the different aircraft systems can be checked using several different cockpit displays. Checking such displays helps the crew verify that the aircraft systems are operating within normal tolerances. These displays are also used to determine the nature of a malfunction, if one occurs. The system displays include alert flags, problem annunciators, and alarm tones for the most serious malfunctions, all of which draw the pilots' attention if a problem occurs. Consequently, the scan of these instruments in the absence of flags or alarms is infrequent.

4.2. RNAV Approach Tasks

This section describes unique aspects of an RNAV approach; that is, only those tasks that are different from those described in Section 4.1.

As previously mentioned, in RNAV the aircraft uses two guidance systems combined with the autopilot and auto-throttles to guide the aircraft between the waypoints programmed into the FMS. The lateral navigation (LNAV) guidance function directs the course of the aircraft, while the vertical navigation (VNAV) guidance function directs the pitch and thrust of the aircraft. When the VNAV mode is engaged and the aircraft is following the vertical path programmed into the FMS, the aircraft is in the VNAV PATH mode.

During descents, as a safety feature, the pilots must specifically pay attention to altitude when using the VNAV PATH mode. As the aircraft descends towards 3D waypoints, it will not keep descending unless the pilots have selected a lower altitude in the MCP, even though the next waypoint's altitude may be below the aircraft's present altitude. That is, the RNAV system (i.e., the FMS) will not automatically change the MCP's target altitude when reaching a waypoint during a descent. To keep descending, the pilots must set the next lower altitude into the MCP before reaching the active waypoint. If a new altitude is not set, the aircraft will establish level flight at the MCP altitude, drop out of the VNAV PATH mode, and enter ALT HOLD mode. For the pilots, this means that during most of the descent and approach, they must stay aware of their altitude and keep setting lower altitudes in the MCP as ATC clears them to those altitudes. When they set a new altitude, the aircraft will maintain the desired descent profile until the next waypoint is reached, provided VNAV PATH is maintained.

However, during an RNAV approach, the pilots must set an MCP missed approach altitude that is higher than the current altitude of the aircraft, even though they intend to keep descending from the FAF to the decision altitude. To properly set the missed approach altitude, the pilots must spin the MCP altitude dial (Figure 7) quickly enough to avoid "capturing" an altitude and dropping out of VNAV PATH mode during the final approach segment, which is fairly time critical and task saturated. Usually, the pilots spin the MCP ALT dial so quickly that they overshoot the desired missed approach altitude, and then dial more slowly down to that desired setting. The alternative is to risk an undesired altitude capture, if they dial the missed approach altitude too slowly.

4.2.1. RNAV Task Descriptions

As with the ILS approach tasks, we make several basic assumptions about the aircraft and set several initial conditions for the beginning of the RNAV approach phase tasks:

- The pilots follow ATC instructions to intercept an RNAV approach procedure and use a chart similar to that shown in Figure 4.
- The approach is programmed into the FMS and both the autopilot and auto-throttles are engaged.
- The PF uses the MCP heading, altitude, and speed functions to comply with ATC instructions until established on the approach, and then uses the approach mode.
- After descending from the FAF altitude, the pilots set the missed approach altitude.
- Upon sighting the runway, the PF assumes manual control of the aircraft, since an automated RNAV approach ends in a missed approach by default.

Again, the reader should note similarities between RNAV and ILS approach tasks. Of course, there are differences, too, as follows.

4.2.1.1. Unique RNAV Tasks

<u>Verify LNAV and VNAV PATH Modes</u>

Verifying that the aircraft systems are in LNAV and VNAV PATH involves looking at the flight mode annunciators on the Primary Flight Display (Figure 8). The green letters "V NAV" and " L NAV" appear in the lower left and lower right corners of the PFD, respectively (as actually depicted in the example PFD of Figure 8). In addition, the pilots verify that the decision altitude waypoint and missed approach path are correctly programmed into the FMS. The ND (Figure 9) shows the waypoints, distance and time to the next waypoint, and indicates the aircraft's actual vertical path relative to the planned path (in the lower right section of the ND in Figure 9).

<u>Check Position along Flight Path</u>

The ND also shows the position of the aircraft relative to the RNAV approach path. This information includes the current heading, the current wind speed and direction, and any relevant traffic or weather.

<u>Set Missed Approach Altitude on the Mode Control Panel</u>

The missed approach (MA) altitude is defined as the altitude to climb to in the event of a missed approach. The MA altitude is on the approach chart (Figure 4, for example has a missed approach altitude of 5000 feet MSL), which both pilots review prior to flying the approach. Setting the MA altitude during an RNAV approach while in VNAV PATH requires spinning the dial fast enough to prevent the aircraft from capturing the new altitude and dropping out of the VNAV PATH mode. After setting the MA altitude, the pilots verify that LNAV and VNAV PATH remain active on the PFD and that the aircraft maintains the desired vertical path.

Executing Missed Approach

If the pilots decide they need to execute a missed approach they perform several actions in quick succession. The sequence involves advancing the throttles to go-around thrust, setting the flaps to 20, establishing a positive climb profile and (usually) retracting the landing gear. Some of these actions will be executed by the automatic systems if the DA is reached during an RNAV approach. A missed approach can be executed during any approach at any point. It is described here because it is more likely to occur during an RNAV approach than an ILS or SVS approach because of a higher decision altitude, which increases the likelihood of not seeing the runway.

There are no unique non-sequential RNAV tasks.

4.3. ILS and RNAV Task Timelines

Tables 3 and 4 list the sequential and non-sequential tasks, respectively, performed by the pilots during the two types of baseline approaches (ILS and RNAV). SVS approach tasks are not yet defined, but many of the approach tasks will be very similar. Because of the variability in the initial approach segment, the task tables start with the intermediate approach segment and finish with main wheel touchdown.

The sequential task table (Table 3) is grouped into sequences of tasks associated with specific events. Usually, the crew will perform a sequence of tasks in response to a location (e.g., nearing the FAF) or communication (e.g., landing clearance) stimulus. Non-sequential tasks are performed throughout an approach, as needed or desired, as opposed to being in response to a specific event. A task execution sequence is usually followed by a period of monitoring as the pilots verify configuration changes or anticipate the next task initiator.

Each flight crew performs the tasks slightly differently. Often the callouts and double-checks occur simultaneously with system setting tasks, especially when a pilot’s task performance timing overlaps with the other pilot’s tasks. As such, an overall time has been given for each sequence of tasks rather than providing timing information for each individual task.

Tables 3and 4 also list the distribution of tasks between PF and PNF in the operator column. Each event is listed with a descriptive title and approximate aircraft position and time remaining to touchdown. Altitudes are in feet AGL; speeds are in knots. Task descriptions are either short statements of an action or, when in quotes, represent a spoken phrase. Types of tasks are discrete, intermittent or continuous. Discrete tasks require a single non-recurrent performance, such as activating or deactivating a system, making a setting, or speaking a phrase. Intermittent tasks require multiple recurrent performances such as monitoring a display. Continuous tasks require variable but uninterrupted performance, such as controlling aircraft heading or speed (McGuire 1991).

For both baseline approach types (ILS and RNAV), the task analysis uses hypothetical, but realistic, examples of approaches into Santa Barbara airport in California. The example RNAV approach chart is Figure 4. We assume that the pilots are following ATC directions during the initial approach segment, rather than the initial segment of the published approach. Such ATC directions are very common and serve to sequence traffic arriving into the airport area from many different directions.

Table 3. B757 approach task timeline for sequential tasks.

Event and Task Descriptions	**Approach Type**	**Operator**	**Task Type**
Receive approach clearance			
• B757 weight = 180000 pounds, visibility is poor. • 3000 feet MSL, ~3000 feet AGL, 200 knots, flaps 5. • ~15 miles from runway threshold. • The ATC communication and read-back take approximately 10 seconds. • In parallel with the PNF read-back, the PF may set the new speed. The task sequence after the read-back takes approximately 10 seconds for the crew to complete. • As the aircraft slows the PF calls for flaps 15.			
ATC Communication: "NASA-113, cleared for the ILS approach to Runway 07, slow to 180."	ILS only	ATC	Discrete
"NASA-113 cleared ILS to 7, slowing to 180"	ILS only	PNF	Discrete
Or: "NASA-113, cleared for the RNAV approach to Runway 33 Left, slow to 180."	RNAV only	ATC	Discrete
"NASA-113 cleared RNAV to 33 Left, slowing to 180"	RNAV only	PNF	Discrete
Check airspeed	ILS & RNAV	PF	Discrete
Set speed to 180	ILS & RNAV	PF	Discrete
Check speed setting	ILS & RNAV	PNF	Discrete
Check speed against reference bugs	ILS & RNAV	PF	Discrete
Set approach mode	ILS & RNAV	PF	Discrete
"Approach mode set"	ILS & RNAV	PF	Discrete
"Flaps 15"	ILS & RNAV	PF	Discrete
Move flaps lever to 15 and state, "Flaps 15"	ILS & RNAV	PNF	Discrete
Aircraft configuration			
• Flap deployment from 5 to 15 degrees takes about 15 seconds. • The aircraft is approximately 13 miles from the runway threshold. • As the flaps lower, the PF monitors attitude, speed, and FMAs.			
Glance at flap indicator to confirm setting	ILS & RNAV	PF & PNF	Discrete
Feel pitch change	ILS & RNAV	PF & PNF	Continuous
Monitor PFD	ILS & RNAV	PF & PNF	Intermittent
Glide slope alive (ILS), or about 11 miles from runway threshold (RNAV)			
• The task sequence listed for this event takes approximately 30 seconds to complete. • Distance from runway decreases to ~ 6 miles (ILS) and to ~10 miles (RNAV).			
"Glide slope alive, gear down"	ILS only	PF	Discrete
"Almost 10 miles from runway, gear down"	RNAV only	PF	Discrete
Lower gear lever and reply, "Gear coming down"	ILS & RNAV	PNF	Discrete
"Flaps 20"	ILS & RNAV	PF	Discrete
Set flaps 20 and reply, "Flaps 20"	ILS & RNAV	PNF	Discrete
Glance at gear lights and state, "Gear down and locked"	ILS & RNAV	PNF	Discrete
Glance at gear lights and reply, "Roger, three green"	ILS & RNAV	PF	Discrete

Glance at flap indicator to confirm setting	ILS & RNAV	PF & PNF	Discrete
"Set speed bug plus 20"	ILS & RNAV	PF	Discrete
Set speed	ILS & RNAV	PNF	Discrete
Check speed setting	ILS & RNAV	PF	Discrete
Monitor PFD	ILS & RNAV	PF & PNF	Intermittent
One dot below glide slope (ILS), or about 9 miles from runway threshold (RNAV)			
• This event begins when the glide slope pointer on the PFD (Figure 7) is next to the dot just above the point where the glide slope is captured (ILS), or within a mile of the RNAV FAF. • The task sequence associated with this event takes 10 seconds to complete.			
"Flaps 25"	ILS & RNAV	PF	Discrete
Set flaps 25 and reply "Flaps 25"	ILS & RNAV	PNF	Discrete
"Set speed bug plus 5"	ILS & RNAV	PF	Discrete
Set speed	ILS & RNAV	PNF	Discrete
Check speed setting	ILS & RNAV	PF	Discrete
Monitor PFD	ILS & RNAV	PF & PNF	Intermittent
Final flaps and complete Before Landing Checklist			
• The task sequence associated with this event takes ~30 seconds to complete.			
"Flaps 30"	ILS & RNAV	PF	Discrete
Set flaps 30 reply "Flaps 30"	ILS & RNAV	PNF	Discrete
"Before Landing Checklist	ILS & RNAV	PF	Discrete
Refer to checklist and read steps	ILS & RNAV	PNF	Discrete
"Gear Down?"	ILS & RNAV	PNF	Discrete
Double-check gear lights; check lever full down	ILS & RNAV	PF & PNF	Discrete
"Down and checked"	ILS & RNAV	PF	Discrete
"Down and checked"	ILS & RNAV	PNF	Discrete
"Flaps 30?"	ILS & RNAV	PNF	Discrete
Check flap indicator setting; check flap lever in detent	ILS & RNAV	PF & PNF	Discrete
"Flaps 30"	ILS & RNAV	PF	Discrete
"Flaps 30"	ILS & RNAV	PNF	Discrete
"Speed brakes armed?"	ILS & RNAV	PNF	Discrete
Check speed brakes	ILS & RNAV	PF & PNF	Discrete
"Armed"	ILS & RNAV	PF	Discrete
"Armed"	ILS & RNAV	PNF	Discrete
"Before Landing Checklist is complete"	ILS & RNAV	PNF	Discrete
"Roger, checklist complete"	ILS & RNAV	PF	Discrete
Glide slope capture (ILS), or FAF (ILS & RNAV)			
• ~130 knots. • ~5 miles from touchdown (ILS) or ~8 miles (RNAV); ~3-4 minutes from touchdown. • The task sequence associated with this event takes about 20 seconds to complete.			
"Glide slope capture"	ILS only	PF	Discrete
"That's the FAF"	RNAV only	PF	
Verify correct FMAs on PFD	ILS & RNAV	PF & PNF	Discrete
Verify thrust and pitch decrease to fly glide slope	ILS only	PF	Discrete

Verify thrust and pitch decrease to fly VNAV path	RNAV only	PF	Discrete
"Setting missed approach altitude" (RNAV only – fast dial spin)	ILS & RNAV	PF	Discrete
Sets missed approach altitude and points to the MCP display	ILS & RNAV	PF	Discrete
"Roger, missed approach altitude set"	ILS & RNAV	PNF	Discrete
Switch radio to tower frequency			
• ~4-7 miles out; ~1500 feet AGL; ~125 knots. • ~2-3 minutes from touchdown. • The time associated with this task sequence can vary depending on how long it takes for the radio dialog. • It should take less than 45 seconds for this whole sequence.			
Scan instruments; if no warning flags, "Flags checked"	ILS & RNAV	PNF	Discrete
"NASA-113 contact the tower; good day"	ILS & RNAV	ATC	Discrete
"NASA-113 switching; good day, sir"	ILS & RNAV	PNF	Discrete
Switch radio to tower frequency	ILS & RNAV	PNF	Discrete
"Santa Barbara Tower, NASA-113, ROCKY inbound"	ILS only	PNF	Discrete
"NASA-113, Santa Barbara Tower, cleared to land Runway 07; winds 060 at 5; ceiling reported at 200 feet"	ILS only	ATC	Discrete
"NASA-113 cleared to land on 7, copy 200-foot ceiling"	ILS only	PNF	Discrete
Or: "Santa Barbara Tower, NASA-113, GOLET inbound"	RNAV only	PNF	Discrete
"NASA-113, Santa Barbara Tower, cleared to land Runway 33L; winds 350 at 5; ceiling reported at 800 feet"	RNAV only	ATC	Discrete
"NASA-113 cleared to land on 33L, copy 800-foot ceiling"	RNAV only	PNF	Discrete
"Cleared to land, ceiling at minimum"	ILS & RNAV	PNF to PF	Discrete
"Roger, cleared to land"	ILS & RNAV	PF	Discrete
Turn on landing lights, if they are not already on	ILS & RNAV	PNF	Discrete
500-foot call-out (ILS) or 1000-foot call-out (RNAV)			
• 500 feet AGL (ILS), or 1000 feet AGL (RNAV). • ~1-2 minutes from touchdown. • (See Transition from Automatic to Manual Flight Event, Table 4.)			
Call 500 feet, speed relative to bug and descent rate	ILS only	PNF	Discrete
Or: Call 1000 feet, speed relative to bug and descent rate	RNAV only	PNF	Discrete
"Roger"	ILS & RNAV	PF	Discrete
100 feet to decision height (ILS), or 100 feet to decision altitude (RNAV)			
• ~300 feet AGL (ILS), or ~800 feet AGL (RNAV). • ~45-90 seconds to touchdown. • (See Transition from Automatic to Manual Flight Event, Table 4.)			
"100 feet to decision height"	ILS only	PNF	Discrete
Or: "100 feet to decision altitude"	RNAV only	PNF	Discrete
"Roger"	ILS & RNAV	PF	Discrete

Decision height (ILS), or decision altitude (RNAV)			
• ~200 feet AGL (ILS), or ~ 700 feet AGL (RNAV); ~30-60 seconds to touchdown. • The point at which the PF begins hand-flying the airplane could take place before this location whenever the runway is sighted (see Transition from Automatic to Manual Flight event in Table 4). • If the runway is not sighted by the time decision height or decision altitude is reached, then the PF initiates a missed approach. • (See Transition from Automatic to Manual Flight Event, Table 4.)			
"Minimums"	ILS & RNAV	PNF	Discrete
If runway is in sight, call out "Runway in sight, landing."	ILS & RNAV	PF	Discrete
Manually fly the landing			
Scanning out window (mostly) and to PFD, airspeed, and vertical speed indicators	ILS & RNAV	PF	Continuous
Manipulate controls as needed for proper touchdown	ILS & RNAV	PF	Continuous
Monitor instruments	ILS & RNAV	PNF	Continuous
100 feet above the runway call-out			
• ~100 feet AGL; ~10 seconds to touchdown.			
"100 feet"	ILS & RNAV	PNF	Discrete
"Roger"	ILS & RNAV	PF	Discrete
Flare and Touchdown			
• ~30 feet AGL; ~117 knots; ~3 seconds to touchdown.			
Flare and let aircraft settle onto main landing gear	ILS & RNAV	PF	Continuous

Table 4. B757 approach task timeline for non-sequential tasks.

Event and Task Descriptions	**Approach Type**	**Operato r**	**Task Type**
Transition from Automatic to Manual Flight			
• This event normally occurs once during the approach and landing phase. The PF decides when to begin manually flying the aircraft. This is usually associated with being able to see the runway. • Once the PF begins flying manually, his or her attention is out the window, primarily, making sure to properly align with the runway. The PNF monitors the instruments and performs the required call-outs.			
Place hands and feet on aircraft controls	ILS & RNAV	PF	Discrete
Press autopilot disengage button when ready to assume control	ILS & RNAV	PF	Discrete
Turn off alarm	ILS & RNAV	PNF	Discrete
Scanning out window (mostly) and to PFD, airspeed, and vertical speed indicators	ILS & RNAV	PF	Continuous
Manipulate controls as needed to maintain proper flight path	ILS & RNAV	PF	Continuous
Monitor instruments to ensure maintaining proper flight path (within acceptable tolerances)	ILS & RNAV	PNF	Continuous

Monitor Flight Path and Progress			
• This task is ongoing throughout all phases of flight and consists of periodic instrument scans. The crew periodically looks at the ND to verify that the aircraft is traveling along its assigned path and at the PFD to verify that the aircraft is at its assigned altitude and appropriate attitude. While the ND and PFD are the primary instrument displays used by the crew, other instruments are occasionally included in periodic scans.			
Monitor PFD	ILS & RNAV	PF & PNF	Intermittent
Monitor ND	ILS & RNAV	PF & PNF	Intermittent
Monitor other aircraft displays and indicators	ILS & RNAV	PF & PNF	Intermittent
Monitor the Radio			
• This task occurs throughout all flight phases, but neither constant nor intermittent attention is required. Instead, attention is directed when a voice is heard from the radio.			
Monitor radio communications	ILS & RNAV	PF & PNF	Intermittent & Discrete
Monitoring Aircraft Systems			
• This task occurs throughout all phases of flight. Even though the aircraft systems are designed to flag or otherwise alert the crew to system problems, the crew will periodically scan the system displays looking for abnormalities.			
Monitor system displays	ILS & RNAV	PF & PNF	Intermittent

4.4. Cognitive Decisions

Now that we have detailed approach tasks, this section discusses a summary of the key decisions during the approach and landing.

4.4.1. Missed Approach

The decision to execute a missed approach could be caused by any of the following:

1. ATC-directed
2. Too high on glide slope
3. Too low on glide slope
4. Too far left or right of the extended runway centerline
5. Too fast
6. Failure to see the runway at decision height because of poor visibility
7. Other weather-related events

4.4.1.1. ATC-Directed

An ATC-directed missed approach is of little interest to this research because it removes the pilots from the decision-making process. The pilots comply when directed to go missed approach, and only have to execute the missed approach as published or as modified by ATC.

4.4.1.2. Too High on Glide Slope

If the aircraft is too high on the glide slope, the aircraft could overshoot the runway. As the aircraft approaches decision height, the decision to continue the landing involves evaluating several quantitative and qualitative factors:

Quantitative factors:

- How high above glide slope is the aircraft?
- Will the aircraft be able to safely touchdown based on it present position?
- How much runway is needed to stop once the airplane touches down?
- How much stopping distance will the runway provide?
  - How long is it?
  - Has the tower directed the pilots to land and hold short of an intersecting runway?
- What are the conditions on the surface of the runway (ice, snow, rain puddles, etc.)?
- Are all needed aircraft systems operational (e.g., brakes, thrust reversers)?
- Is there enough fuel for a 20-minute missed approach and another approach?
- How heavy is the aircraft? (A light airplane stops more quickly than a heavy one.)

Qualitative factors that may affect the decision to continue to a landing:

- Maintaining schedule
- Passengers with missed connections
- The pilots want to get off this airplane and go home

With a strong emphasis on safety, if any of the quantitative factors clearly indicate that a landing might be unsafe, the pilots execute a missed approach. In addition, the pilots do not need to wait until the decision point to do so – they can execute a missed approach at any time during an approach. If the quantitative factors do not make the decision easy, one would expect that the qualitative factors might become relevant.

4.4.1.3. Too Low on Glide Slope

If the aircraft is too low on the glide slope, the aircraft could hit the ground prior to the runway threshold. The decision to continue the landing involves the same qualitative factors as the "too high on glide slope" case, but there are fewer quantitative factors:

Quantitative factors:

- How far below glide slope is the aircraft?
- Are there obstacles or high terrain that may jeopardize safety?
- If the aircraft continues the descent at the same rate, where will it touchdown?
- How much of a descent rate adjustment is needed to fly *safely* over the runway threshold?
- Is there enough fuel for a 20-minute missed approach and another approach?

4.4.1.4. Too Far Left or Right of Extended Runway Centerline

If the aircraft is too far left or right of the runway, the aircraft might not physically be able to reach the runway. Or it might land on the runway, but then skid or roll off of it. The decision to continue the landing involves similar qualitative factors as the other cases, but fewer quantitative:

- Can the aircraft align with the runway centerline in the amount of time left until touchdown?
- Is there enough fuel for a 20-minute missed approach and another approach?

4.4.1.5. Too Fast

The kinetic energy of an aircraft is proportional to its speed squared. Hence, if a pilot is coming in 20% faster than desired, the braking energy to stop the aircraft will be 40% higher than planned. Because of this energy dissipation issue, pilots are very concerned with their speed, particularly at airports with short runways.

- How much runway is needed to stop, once the aircraft touches down?
- How much will the runway provide?
  - How long is it?
  - Has the tower directed the pilots to land and hold short of an intersecting runway?
- What are the conditions on the surface of the runway?
- Are all needed aircraft systems operational (e.g., brakes, thrust reversers)?
- Is there enough fuel for a 20-minute missed approach and another approach?
- How heavy is the aircraft? (A light airplane stops more quickly than a heavy one.)

4.4.1.6. Failure to See the Runway at the Decision Point Because of Poor Visibility

If the runway is sighted and the aircraft is not properly aligned (too high, too low, etc.), one of the other decisions listed above becomes relevant. Until the runway is sighted though, the following elements are considered by the pilot:

- The pilots have already received a report of the cloud ceiling altitude.
- As the aircraft approaches the ceiling altitude, the pilots have an expectation of seeing the clouds begin to break up.
- Below the ceiling altitude, but above decision height:
  - If there is no indication that the clouds are breaking up, the pilots mentally prepare for a missed approach.
  - If the clouds are starting to break up, the pilots plan to continue with the approach.
- Just prior to decision height:
  - The pilots are looking for any changes in visibility to reaffirm or disprove their earlier predisposition to:
    - Continue the landing (most likely the runway is in sight by this time).
    - Execute a missed approach (most likely the runway will not appear).
- At decision height, unless something very unusual happens (e.g., the aircraft exits a very well-defined cloud layer instantaneously), the pilots have already made their decision and act accordingly.

4.4.1.7. Other Weather-Related Events

Wind shear and micro-bursts are other weather related events that can cause missed approaches. Pilots know these conditions are out of their control so they are likely to be very conservative with their decision to continue the approach. Pilots understand that any missed approach, regardless of the circumstances, will not reflect negatively on their skills, abilities or job security.

4.4.2. Controlling Speed

Controlling speed is not a single decision point *per se* because the need to control speed is often required throughout the approach. Controlling speed is a more difficult task for the B757 during approach compared to older commercial aircraft (e.g., B727) because of its very good glide performance (high lift over drag ratio). While the auto-throttles can automatically increase thrust if additional speed is needed, there is no *automatic* equivalent for slowing the B757. For quick deceleration, the good glide performance precludes just setting the engines to idle, particularly when the aircraft is descending anyway. The speed brakes can be used manually to slow down the aircraft, but it violates procedures to use the speed brakes during the final approach segment. The flaps provide additional drag, but flap extension is a function of airspeed; they cannot be lowered at too high of a speed or else they will be damaged. Another method for slowing the aircraft is to extend the landing gear earlier in the approach phase since the landing gear adds considerable drag to the aircraft. But, the wheels also have a maximum speed at which they can be extended. The point is that the pilots must focus on reducing speed as early in the approach as possible.

Sometimes, though, this is easier said than done. For example, during the initial approach segment when ATC has several aircraft sequenced for landing, if all but the first two aircraft are properly spaced, but the distance between the first and second is too close, then it is much easier for ATC to instruct the first aircraft to speed up to increase its spacing from the second aircraft. This saves the controller from having to instruct the several other aircraft further back to slow down. Therefore, even though the pilots want to slow down early in the approach, it is not always possible.

Approaches are fraught with the potential for problems and errors, as alluded to above. Next, we further amplify potential approach problems and errors.

4.5. Approach Problems and Errors

Many aviation accidents and incidents are due to chains or sequences of problems. Some of the problems that are part of these chains are known as *latent* errors that were committed or occurred either well before or early in the flight and compound or create problems later on in the flight. Although these errors and error chains are important in terms of aviation safety, the focus of this research has been on the approach and landing phases of flight. In addition, the human performance modeling efforts that follow this task analysis research will focus on comparisons of equipment used primarily during approach and landing. As such, the errors discussed in this section are limited to those that can occur during the approach phase, but are not specific to either ILS or RNAV approaches.

<u>Aircraft Spacing</u>

Spacing errors become a problem as ATC tries to prescribe aircraft locations and require maneuvers that may be difficult or impossible for the crew to perform. Problems can occur when ATC asks the crew to maintain a particular speed when they really need to be slowing or when they are asked to maintain spacing behind an aircraft they know can slow down faster than the B757. It is up to the crew to keep out of bad situations. The pilots use TCAS and radio call information to maintain their awareness and spacing from other aircraft, while still complying with ATC instructions. A result of a spacing error can be insufficient time for a preceding aircraft to clear the runway before the following aircraft lands. In this case, the following aircraft may have to execute a missed approach.

<u>Stabilization Gates</u>

The crew must be able to achieve particular stabilization gates. That is, at certain locations during the approach, the crew must have achieved a certain attitude and speed in order to continue the approach. At 1000 and 500 feet AGL the speed, sink rate, and alignment suitable for landing should be within tolerances. If the aircraft is not stable then a missed approach is mandatory.

<u>Speed Brakes</u>

For each of the last three problem areas there is some help on the B757. If the pilots find that they are coming in too high or fast, the PF may use the speed brakes. On the B757, using speed brakes is part of the standard procedure during approach and landing. The PF normally deploys the speed brakes for a short period to help the autopilot achieve a necessary altitude or speed requirement. While the use of the speed brakes is common practice for the B757, there are also problems associated with using the speed brakes. The primary error is to forget that they are deployed and try to land. Doing so induces a high sink rate and increased deceleration that the autopilot and auto-throttle will try to counteract. In addition, if the speed brakes are deployed during landing, the tail is likely to strike the runway upon flaring.

FMS Reprogramming

Another problem relates to the reprogramming of the FMS. Due to weather or traffic issues, ATC may change the approach clearance from that expected by the pilots and programmed into their FMS. Such changes can occur at any point during the flight and can be a common occurrence at crowded airports. If such a change is made, the pilots may want to reprogram the FMS to reflect the new flight plan, but may have insufficient time to do so.

The general rule is that pilots should not attempt to reprogram an approach when below 10000 feet MSL. The problem is that there may not be enough time to make the changes and still perform the necessary slowing and configuration tasks. In addition, the attention of the PNF working with the FMS is directed down and away from other instruments and the cockpit windows. This attention mismanagement can yield failures to monitor altitude restrictions or course changes, both of which can have disastrous results.

Radio Frequency

Approach frequency errors of the communication and navigation radios can also occur. This happens usually as a result of haste, the failure to double-check, or through simple entry errors. Such errors are usually caught using sufficient crew coordination practices. Also, if two navigation radios are set to mismatched frequencies then warning flags alert the crew. However if both pilots fail to set the radios or make appropriate changes, there are no flags to alert them to the problem. Unlike the navigation radios, there are no alarms for miss-setting the communication radios. Usually, upon changing the radio frequency, the PNF initiates a call or hears other communication. If no communication is heard or there is no response, the PNF normally returns to the previous frequency to request a repeat of the new frequency.

The consequences of errors setting the communication radios can range from not receiving a clearance, thus incurring a missed approach, to more serious issues of spacing in heavy traffic patterns. The consequence of errors setting the navigation radio frequency can be a failure to capture the localizer, which can lead to serious position errors.

Distractions

Distractions are not uncommon during flight. There are many circumstances that can divert the crew's attention from a current task. Some of these are events or issues that must be attended to while others can represent simple nuisances. If the pilots become distracted, they may not remember to return to an important task or may not complete that task in a timely manner.

Other distractions may function as performance shaping factors that make normal tasks more difficult. Changes made by ATC to the approach plan can become distractions. This is especially true if the aircraft is close to the airport and the crew has to change routing during an already busy phase of flight. A high volume of communication traffic on the radio can distract the crew from other tasks as they attempt to comprehend all the information that is being spoken. Periods of high air traffic associated with the approach phase at busy airports will also provide distractions as the crew attempts to maintain visual contact and spacing from nearby aircraft. Weather can actually be more distracting when it is minor. Serious weather, that significantly reduces visibility, usually results in changes to the aircraft spacing rules and runways that are used. However, scattered or intermittent clouds may make it difficult to identify and follow other aircraft during marginal visual approach conditions. Likewise, approaches and landings done at night over large brightly lit areas can make it difficult to see the lights of other aircraft and the airport as those identifying lights become lost in background glare. Finally, equipment problems or failures can represent serious distractions depending on the system, severity of the problem, and phase of flight.

4.5.1. RNAV Specific Errors

The following errors are specific to RNAV approaches. In each case they involve undesirable changes in the automatic flight modes based on actions or inactions by the crew.

Lateral Navigation

The LNAV mode keeps the aircraft on the lateral course to the next waypoint programmed in the FMS. The FMS accounts for any required heading changes to reach the active waypoint. Small heading changes might be required due to crosswinds. If the aircraft is in the Heading Select mode rather than LNAV, the aircraft will follow the set heading without accounting for any crosswinds. In this situation, the aircraft might be blown off the proper course to the next waypoint. If VNAV is engaged with Heading Select, then the vertical profile of the approach would continue the descent, while the lateral course might be erroneous. There are no system warnings to prevent or recover from this error; it is up to the pilots to detect and correct, which is especially problematic when visibility is poor. Certainly the pilots would notice lateral alignment problems when the runway is sighted. However, if this error occurs in poor visibility during approach to an airport with nearby steep terrain, then the result could be disastrous.

Failure to Set Decision Altitude (DA)

Prior to reaching the FAF the pilots must set the DA in the MCP. In the example from Figure 4, the aircraft is flying level at 1800 feet approaching the FAF at GOLET. If they do not set the DA in the MCP prior to reaching GOLET, the aircraft systems will switch to Altitude Hold mode and drop out of VNAV PATH. Although LVAV will continue to provide course guidance, the aircraft will continue to fly at 1800 feet. The cockpit systems provide some help to prevent this situation by alerting the pilots to reset the MCP altitude just prior to reaching a waypoint where a change in altitude is programmed. This alert message is presented on the front instrument panel (below the engine instruments in Figure 6) but does not include an aural annunciation. Of course, the pilots might notice that the descent has not begun as expected after passing through the waypoint.

Setting Wrong Decision Altitude

If the pilots do set a DA prior to reaching the FAF, it is still possible to set an incorrect altitude into the MCP. If they set an altitude that is higher than the desired DA, the aircraft will execute a missed approach sooner than expected. If the crew is able to quickly diagnose the problem they will then have to decide whether or not to continue with the missed approach or take manual control of the aircraft and try to land. If the pilots set an MCP altitude that is lower than the desired DA and they do not break-out below the cloud deck prior to reaching that altitude, the aircraft will execute the missed approach but at a lower altitude than the published DA. The consequence of this error depends upon how close the aircraft flies to the terrain and whether or not the PF takes manual control before the miss-set DA.

Failure to Set Missed Approach (MA) Altitude

Once the aircraft is stabilized on the final descent above the DA, the crew needs to set the missed approach altitude in the MCP so that the autopilot executes the correct missed approach at the DA. If the crew fails to set the MA altitude, the aircraft will switch to Altitude Hold mode at the DA. In the example RNAV approach (Figure 4), Alt Hold would occur at 700 feet (even though the DA is 650 feet MSL, the MCP can only be set to the nearest hundred feet). The aircraft would start the missed approach by accelerating and turning left toward GOLET, but it would level at 700 feet rather than climbing to 5000 feet. At first, to the pilots, the aircraft would feel like it was performing the missed approach properly; the aircraft would stop descending and the throttles would move forward. However, the aircraft would not begin climbing. The problem from the crew's perspective is one of timely recognition, especially with obstacles and high terrain nearby. The pilots expect the aircraft to halt the descent and accelerate, and at first these expectations are not violated. They might even feel as though they are climbing as they accelerate. To recognize the problem in time when in the clouds, the pilots would have to carefully observe the vertical speed indicator, altimeters, and PFD pitch angle.

Altitude Capture while Setting Missed Approach Altitude

While spinning the MCP altitude dial to set the missed approach altitude, it is possible that the aircraft could enter Altitude Capture mode. That is, if there is a pause or slowing while spinning the dial, the aircraft may capture an altitude between the DA and the MA altitude settings, which is not what the pilots intend. In the RNAV approach example (Figure 4), while the aircraft is descending from 1800 feet and the crew tries to reset the MCP from the 700-foot DA to the 5000-foot missed approach altitude, the aircraft might capture an altitude of 1500, for example. If this occurs, the autopilot will switch to Altitude Hold mode upon reaching 1500 feet MSL and fly level at that altitude. The flight mode annunciators will switch to Alt Capture then to Alt Hold and the altitude hold light on the MCP will illuminate. Per procedures, if VNAV PATH is lost during this final descent and if the pilots do not see the runway, they must execute a missed approach.

4.6. Information Requirements and Situational Awareness

Naturally, professional pilots seek to avoid problems and errors. Therefore, they spend a large portion of their time during flight maintaining and updating an accurate mental picture of where the aircraft is, how the flight is progressing relative to the plan, and predicting how changes will affect the flight plan in the future. It is this situation awareness that forms the basis for decision-making during flight. The airline industry recognizes that pilot SA is an important component of flight safety, and has expended great effort to train and teach SA and decision making skills as a part of introductory and recurrent pilot training. In addition, a number of studies have focused on SA and the information requirements of commercial pilots during flight. We examined three such studies in reference to the baseline and SVS approach task analyses.

The first two studies (Ververs, 1998 and Schvaneveldt, 2000) both focus on the relative importance of information during different phases of flight. In both cases, pilots were provided with a survey that included a list of common information available during flight. The subject pilots ranked the relative importance of each piece of information across different phases of flight. Both studies included the approach and landing phases and both studies presented similar results in terms of what pilots thought were the most important cues.

The third study (Endsley, 1998) created an exhaustive list of every type of information desired by pilots to generate complete and accurate SA throughout an entire flight. The information requirements were based on the goals and decisions that pilots make throughout each phase of flight. The result is a highly detailed list of SA information requirements for each of the three levels of SA – perception, comprehension and projection – included in Endsley's taxonomy of situational awareness. She generated this list to help future cockpit designers, and noted that some of the information desired by pilots is not currently provided by cockpit systems. Also, the study recognized that some SA information elements are more important than others during different phases of flight. However, Endsley did not prioritize this information as was done in the other two studies.

The combined knowledge from these three studies indicates that SVS has the potential to improve pilot SA by presenting information that is not currently displayed in commercial aircraft, and by emphasizing information that pilots judge to be most important, especially during approaches.

4.7. SVS Approach Benefits

The use of SVS is mainly notional at this point since there are no FAA-approved procedures for such systems. The concept of operations information (Section 3.3.2.1) provides ideas and goals for what the system will do and how it should be designed, but does not discuss how pilots will actually use it down at the level of individual task steps. Likewise, simulations and test flights using SVS have focused on specific research issues related to the design or use of the system, independent of other flight concerns. Given these limitations, we have not attempted a task decomposition for SVS approaches. We do, however, discuss some tasks that probably will not change with SVS use. Following that, we discuss how SVS may change the cues that are available to pilots.

The procedures for aircraft stabilization during approach are unlikely to change due to the use of SVS. The speeds and flap settings used to slow and stabilize the aircraft, the required callouts between pilots, the use of checklists to verify the completion of required tasks, and radio calls to ATC will all still occur during the approach sequence. One change that may occur relates to the timing of the configuration steps. In poor visibility conditions, configuration steps are usually completed earlier in an approach than during good visibility conditions. The use of SVS may allow pilots to delay some configuration steps during poor visibility, as if they were flying in good visibility conditions. Delaying some configuration steps saves fuel.

4.7.1. Situation Awareness with and without SVS

This section focuses on the differences in cues available from SVS versus the cues available in the standard B757 instrumentation during poor visibility conditions. Since the most significant benefit of SVS touted by an SVS test pilot was the improved SA, it is important to compare the cues from SVS with standard B757 instrumentation to identify how SVS impacts SA. To give this discussion more context, though, it is worthwhile to first present Endsley's (1999) decomposition of the three levels of SA:

- **Level 1 SA – Perception of the elements in the environment.** The first step in achieving SA involves perceiving the status, attributes, and dynamics of relevant elements in the environment. The pilots need to accurately perceive information about their aircraft and its systems (airspeed, position, altitude, route, direction of flight, etc.), as well as the weather, air traffic control (ATC) clearances, emergency information, and other pertinent elements.
- **Level 2 SA – Comprehension of the current situation.** Comprehension of the situation is based on a synthesis of disjointed Level 1 elements. Level 2 SA goes beyond simply being aware of the elements that are present to include an understanding of the significance of those elements in light of the pilots' goals. Based upon knowledge of Level 1 elements, particularly when put together to form patterns with the other elements, a holistic picture of the environment is formed, including a comprehension of the significance of information and events. The pilots need to mentally combine disparate bits of data to determine, for example, the impact of a system malfunction on another system, or deviations in aircraft state from expected or allowable values. Novice pilots might be capable of achieving the same Level 1

SA as more experienced pilots, but may not integrate various data elements along with pertinent goals to comprehend the situation as well as more expert pilots.

- **Level 3 SA – Projection of future status.** It is the ability to project the future actions of the elements in the environment – at least in the near term – that forms the third and highest level of situation awareness. This level is achieved through knowledge of the status and dynamics of the elements and a comprehension of the situation (both Level 1 and Level 2 SA). For example, pilots must not only comprehend that a thunderstorm – given its position, movement and intensity – is likely to create a hazardous situation within a certain time period, but they must also determine what airspace will be available for route diversions, must ascertain where other potential conflicts may develop, must plan appropriate path changes, and must request changes from ATC. This level of SA enables pilots to decide on the most favorable course of action.

When comparing SVS to standard B757 instrumentation, the above descriptions of the three levels of SA highlights a very significant strength of SVS; that is, SVS provides Level 2 and 3 SA directly to the pilots. For example, consider a missed approach at an airport surrounded by steep terrain (Figure 4). If the velocity vector is above the terrain profile on the SVS display, the pilots know that the projected location of the aircraft will ensure flying clear of the terrain (assuming, of course, that the aircraft performance does not change significantly, for example, due to wind shear or a loss of thrust). Even if the missed approach is performed by the autopilot and auto-throttles rather than manually, the pilots gain confidence in the automation and missed approach execution when they see the velocity vector rise above the terrain.

In contrast, for a non-SVS-equipped B757, the pilots would have to integrate Level 1 SA from different sources to have a similar level of confidence. The elevation of the terrain comes from the approach charts; aircraft heading comes from the navigation display; vertical speed comes from the vertical speed indicator; and, distance to the terrain comes from mentally overlaying the terrain information on the navigation display. Granted one of the primary reasons for requiring pilots to follow missed approach procedures is so that they do *not* have to integrate all this information to ensure terrain clearance. But the point here is if the pilots do not precisely follow the MA procedure or if the aircraft is not where the pilots think it is (for whatever reason – distractions from cockpit alerts, ATC communication, etc.), then ensuring terrain clearance becomes workload intensive, exacerbating an already unusual situation. With SVS, terrain awareness is quickly acquired from a single display. Furthermore, the feedback provided by the velocity vector relative to the terrain image gives a prediction of what is needed (in terms of control inputs) for terrain clearance. In contrast, an awareness of terrain clearance is very difficult to achieve with current B757 instrumentation.

SVS also provides the pilots with the ability to crosscheck displays that receive data from different sources, therefore increasing confidence in total system accuracy, or revealing a malfunction sooner than might otherwise be possible. The SVS display uses GPS position to present the aircraft's position relative to the terrain. On the other hand, runways equipped with ILS transmitters have a localizer and glide slope that are independent of GPS. If the pilots are flying an ILS approach in poor visibility, they can crosscheck the PFD's ILS guidance with the SVS display. If the PFD's flight director is aligned with the localizer and glide slope, then the pilots should see that the SVS display also shows proper runway alignment. If both displays agree, then the pilots have added confidence in the information. If the displays disagree, the pilots can execute a missed approach and then investigate the source of the problem in a more controlled and safer environment. Pilots routinely crosscheck their altitude readings among independent altimeters, for example. SVS enables crosschecking other critical flight information, as well.

Even though it seems obvious that SVS should enhance pilots SA, Stark (2001) reported that the use of tunnel navigation did not increase the pilots' situation awareness during the approach phase as was expected. Stark suggests that the ease of using the tunnel navigation to maintain the approach path may somehow reduce the amount of information the pilots gather as they fly an approach. It is also possible that the environment of the test, which limited the tasks and decisions required of the pilots, contributed to a lowered sense of, or need for, SA. Stark's results suggest that it is worth further investigating pilot interactions with SVS and its impact on SA.

4.7.2. Pilot Interaction with SVS

An interview with one of the few pilots who has flown tests with an SVS compared the use of SVS to what many pilots experience when using a "hood" with an instructor pilot during instrument procedure practice. The hood blocks a pilot's view out the window, forcing him or her to rely solely on the flight instruments – an excellent practice method. However, hoods are not perfect vision obstructions, and so pilots will sometimes peek around the hood out the window, especially during a difficult approach. SVS provides a constant "peek out the window," according to the SVS test pilot. The idea is that no matter how much pilots trust their instruments, they are aware that interpreting the instrument indications has the potential for error, and that a "peek at the runway" provides an immeasurable sense of confidence. The SVS test pilot described his interaction with SVS in such phrases as, "increased comfort level," "reduced worry," and "increased confidence in the approach."

Given these statements, it is tempting to characterize the use of SVS as the same as flying during good visibility daylight conditions. However, the SVS test pilot rejected this notion indicating that the cues provided by SVS create an interaction that falls somewhere between scans of the instruments and the out-the-window view during good visibility. With this understanding in mind, we next define the differences between the cues provided by SVS and what a pilot sees or does during good visibility daylight conditions. The following section focuses on those differences. We find that the SVS cues are highly dependent on the implementation of the system and, as such, use variations in the NASA and Rockwell Collins concepts (as shown in Figure 11) to present the pilot-SVS interaction.

4.7.2.1. Closure and Crossing Rate Interpretation

Pilots have learned to interpret the position and alignment of the aircraft as a function of the cues provided by the change in the view of the terrain outside the cockpit. Closure and crossing rates are interpreted by the perceived rate of change of the size of features ahead of, and around, the aircraft. The SVS terrain map provides closure and crossing rate cues similar to those available during good visibility daylight conditions. However, the value of the cues seems to decrease with the level of <u>fidelity</u> of the view of the terrain. The SVS test pilot indicated that while the generic texturing of the Rockwell Collins concept provided more of such cues than did ILS instrumentation, the detail provided by the photo texturing of the NASA concept was easier to interpret compared to the Rockwell Collins concept.

There are also two issues related to the <u>field of view</u> provided by SVS. The cockpit windows provide a much greater usable field of view than is provided by the SVS display (Figure 12). Being able to see terrain features to the side of the aircraft during good visibility daylight conditions increases the available cues used for closing and crossing rate interpretation. The SVS concepts limit the pilot to the terrain cues that are ahead of the aircraft. This limitation is particularly pronounced while turning the aircraft. During good visibility daylight conditions, pilots look out a side window or move their heads forward to obtain terrain cues during a turn. The current SVS display does not support looking into a turn the way the cockpit windows do.

Also, in a crosswind, pilots orient the nose of the aircraft to counteract the effect of the crosswind. The result is that the nose of the aircraft is angled relative to the actual direction of travel. In this situation pilots do not look over the aircraft nose, rather they look in the direction of travel by turning their heads in that direction. In this way the closure and crossing rate interpretation remains stable relative to the terrain and direction of travel. The SVS display limits the terrain view to the direction that the aircraft nose is pointing. As such, the perception of closure and crossing rates can be disrupted in a crosswind situation.

Instrument approach procedures are designed to aid pilots during poor visibility and to help them fly the aircraft below the cloud deck to a point where visual sighting of the runway and a normal landing are possible. The SVS test pilot indicated that when first breaking out below the clouds there is an adjustment period between using the ILS instruments and becoming oriented based on the visual out-the-window scene. The difficulty making this adjustment may relate to identifying terrain features and correlating those features to the aircraft position. He indicated that using SVS might aid that transition as the pilots should have a better idea of the terrain features and their orientation prior to breaking out below the clouds.

4.7.2.2. Tunnel Navigation

The tunnel navigation component of the SVS display (Figure 11) provides navigation cues beyond what are available to pilots during good visibility daylight conditions. The tunnel can be thought of as a combination of the glide slope and localizer guidance cues because the bottom, top and sides of the tunnel help the pilots to keep the aircraft within a defined approach corridor. However, the tunnel exceeds localizer and glide slope guidance because it presents approaches with altitude steps, or circling approaches that require the pilot to align with the runway after an arcing and descending turn. The tunnel allows pilots to follow a different visual representation of the approach rather than relying on their interpretation of separate lateral and vertical path cues to maintain the approach profile. The lateral and vertical guidance of ILS and RNAV approaches are each 2D representations of the approach. Whereas the tunnel is more of a 4D representation in that it gives a 3D combination of lateral and vertical guidance, plus a depiction of the desired path into the future (3D + time = 4D).

The different graphical implementations of the tunnel represent different levels of cue availability. The Rockwell Collins tunnel shows a continuous connected path that may make it easier to see the extended tunnel ahead of the aircraft allowing the pilots to anticipate changes in the flight path. The NASA implementation does not enclose the tunnel or connect the individual tunnel pieces. The NASA concept may make it more difficult to see the tunnel as it extends ahead of the aircraft, thus limiting the ability of the pilots to anticipate flight path changes. However, there is a trade-off in terms of screen clutter. The Rockwell Collins tunnel uses more graphics overlaid on the terrain image. This clutter could, for example, make it difficult to distinguish between the line for the artificial horizon and lines from the tunnel, especially during turns, thus making undesirable attitude changes more likely.

Both SVS concepts use an aid to help the pilots maintain position within the tunnel. The NASA concept uses the ghost aircraft to lead the pilots along the path. The task involves keeping the velocity vector aligned with the tail of the ghost aircraft. The Rockwell Collins concept uses a magenta colored box that borders the tunnel and leads the aircraft along it. The pilots keep the velocity vector within the magenta box as it travels ahead of the aircraft. While the idea of traveling along the tunnel is highly intuitive and the idea of following something along it makes the task easier, there are some issues related to cue differences and performance between the magenta box and ghost aircraft implementations. The magenta box of the Rockwell Collins concept predicts the aircraft position along the tunnel 5 seconds into the future (if the pilots make no control changes). If the velocity vector is within the box, the aircraft will fly within the box position 5 seconds from the current time. However, the task of positioning the velocity vector on the ghost aircraft in the NASA concept may require the PF to make constant small control changes because the size of the ghost is relatively small compared with the movements of the velocity vector. In contrast, the task of keeping the velocity vector within the Rockwell Collins' magenta box requires a lower level of workload because the box presents a much greater target area on the SVS screen, which creates the impression that the PF has larger maneuver tolerances than with the ghost aircraft. The SVS test pilot indicated that the difference in workload was not so much with control manipulations as it was in the attention required to perform the task.

Another issue relates to convention violations within the representations. The first concerns the predictive nature of the magenta box versus the standard use of the velocity vector. The velocity vector has traditionally represented the flight path of the aircraft based on control inputs at the current time. The task of positioning it within the magenta box that represents a position 5 seconds into the future violates that convention. The second issue involves the implementation of the ghost aircraft. Pilots with military formation flying experience (many commercial pilots have military backgrounds), who follow an aircraft, know that they need to turn inside a lead aircraft's turn to maintain a fixed distance from that lead aircraft. While the ghost aircraft resembles such a lead aircraft, it violates the convention in two ways. First, the ghost aircraft cues a turn by yawing rather than banking, which makes it difficult to notice when the turn begins. Second, the task of following the ghost aircraft involves keeping the velocity vector on the ghost, rather than ahead of it during a turn, as formation flying practices dictate. It is impossible to know what the effects of these convention violations will be, but it is anticipated that, with training and increased use, the differences from convention may not present insurmountable problems.

5. Future Work

Human performance modeling leverages systems engineering resources for new systems by identifying issues early in the design and development stages (while changes can still be made relatively easily, compared to engineering changes after fielding a system). HPM also reveals human-system integration issues that otherwise might not be discovered until prototype testing or usability analyses. While it seems logical that SVS will help pilots fly more precise approaches, with better SA and fewer errors, this is a testable hypothesis. The next steps in the ongoing NASA AvSP HPM research are to build pilot performance models and to begin testing the hypothesis. Modeling teams supporting the Human Performance Modeling Element will next develop pilot performance models based upon the task analyses and run digital experiments looking for pilot performance, workload, or related issues to show the tangible benefits of SVS compared to baseline conditions.

6. Suggested Reading

Fundamentals of Air Traffic Control by M. Nolan. Provides an excellent description of ILS, instrument approach procedures, runway lighting, and ATC communication phraseology.

An Exploration of Function Analysis and Function Allocation in the Commercial Flight Domain, by J. McGuire et al. of Douglas Aircraft (over 300 pages). Provides a very detailed functional analysis for a flight from LA to NY. Excellent source for event timeline information.

Key Cognitive Issues in the Design of Electronic Displays of Instrument Approach Procedure Charts by Monterey Technologies. The main document doesn't apply specifically to HPM of the approach, but it has a very interesting 34-page appendix, which actually includes a Conceptual Graph Structure of the ILS approach. Also contains a high-level task analysis.

The Boeing 757/767 Simulator and Checkride Procedures Manual by M. Ray. This unofficial manual provides excellent insight into recovering from off-nominal events or conditions, and highlights the key items to remember to stay out of trouble in the first place.

The Pilot's Guide to the Modern Airline Cockpit by S. Casner. This "technical, but doesn't read that way" book very clearly explains the new generation flight deck with an emphasis on the FMS and guidance modes.

Situation Awareness Requirements for Commercial Airline Pilots by M. Endsley et al. This paper breaks SA requirements down by phase of flight.

Priority and Organization of Information Accessed by Pilots in Various Phases of Flight by Schvaneveldt et al. Provides insight into what information is most important to pilots, decomposed by flight phase.

Understanding a Pilot's Tasks by P. Ververs. Similar to the above, with a slightly different emphasis.

NASA's Aviation Safety Program web site (http://avsp.larc.nasa.gov/images_svs.html) contains information related to the entire SVS project, including history, concept descriptions, and simulator and flight test documentation.

Flight Test Evaluation of Tactical Synthetic Vision Display Concepts in a Terrain-Challenged Operating Environment by Bailey et al. This is the first report documenting the results of the Eagle County Regional Airport flight test. It includes results relating to terrain awareness, workload, and SA.

Preliminary Examinations of Situational Awareness and Pilot Performance in a Synthetic Vision Environment by Stark et al. This paper reports results of an early SVS simulator-based test.

7. References

Alter, K.W. & Regal, D.M. (1992). *Definition of the 2005 Flight Deck Environment* (NASA Contractor Report 4479 for NASA contract NAS1-18027. Hampton, VA: NASA Langley Research Center.

Bailey, R.E., Parrish, R.V., Arthur, J.J., & Norman, R.M. (2002). *Flight Test Evaluation of Tactical Synthetic Vision Display Concepts in a Terrain-Challenged Operating Environment*. Hampton, VA: NASA Langley Research Center.

Casner, S. (2001). *The Pilot's Guide to the Modern Airline Cockpit*. Ames, IA: Iowa State University Press.

Clay, M. (1993). *Key Cognitive Issues in the Design of Electronic Displays of Instrument Approach Procedure Charts* (DOT/FAA/RD-93/39. NTIS #PB94-141991). Monterey, CA: Monterey Technologies.

Degani, A., Mitchell, C.M., & Chappel, A.R. (1995). Task Models to Guide Analysis: Use of the Operator Function Model to Represent Mode Transition. *Proceedings of the Eighth International Symposium on Aviation Psychology*. Columbus, OH: The Ohio State University.

Endsley, M.R. (1999). Situation Awareness and Human Error: Designing to Support Human Performance. *Proceedings of the High Consequence Systems Surety Conference*. Albuquerque, NM: Sandia National Laboratory.

Endsley, M.R. (2000). *Evaluation of Situational Awareness in Flight Operations Employing Synthetic Vision Systems* (SATECH-00-11). Marietta, GA: SA Technologies, Inc.

Endsley, M.R., Farley, T.C., Jones, W.M., Midkiff, A.H., & Hansman, R.J. (1998). *Situation awareness requirements for commercial airline pilots*. Cambridge, MA: Massachusetts Institute of Technology's International Center for Air Transportation.

Hart, S., & Staveland, L. (1988). Development of the NASA-TLX: Results of empirical and theoretical research. In P. Hancock and N. Meshkati (eds.), *Human mental workload*. Amsterdam: North-Holland.

Leiden, K., Keller, J.W., & French, J. (2002). *Information to Support the Human Performance Modeling of a B757 Flight Crew during Approach and Landing* (NASA Ames contract NAS2-99091. Boulder, CO: Micro Analysis & Design.

McGuire, J.C., Zich, J.A., Goins, R.T., Erickson, J.B., Dwyer, J.P., Cody, W.J., & Rouse, W.B. (1991). *An Exploration of Function Analysis and Function Allocation in the Commercial Flight Domain* (NASA Contractor Report 4374 for NASA contract NAS1-18028). Hampton, VA: NASA Langley Research Center.

NASA (2001). *Synthetic Vision Fact Sheet* (FS-2001-08-48-LaRC). Hampton, VA: NASA Langley Research Center.

Nolan, M. (1994). *Fundamentals of Air Traffic Control*. Belmont, CA: Wadsworth Publishing Company.

Norman, R.M. (2001). *Synthetic Vision Systems (SVS) Description of Candidate Concepts* (NAS1-20342), from http://avsp.larc.nasa.gov/images_svs.html.

Norman, R.M. (2002). *Synthetic Vision Systems (SVS): Description of Candidate Concepts* (NAS1-00106). Hampton, VA: NASA Langley Research Center.

Ray, M. (2001). *The Boeing 757/767 Simulator and Checkride Procedures Manual*. Temecula, CA: University of Temecula Press.

Schvaneveldt, R.W., Beringer, D.B, & Lamonica, J.A. (2000). Priority and Organization of Information Accessed by Pilots in Various Phases of Flight. *The International Journal of Aviation Psychology, 11(3)*, 253-280.

Small, R.L. & Bass, E.J. (2000). Certify for success: A methodology for human-centered certification of advanced aviation systems. In J.A. Wise, & V.D. Hopkin (eds.), *Human Factors in Certification*. Mahwah, NJ: Lawrence Earlbaum Associates, pp. 139-150 (ISBN 0-8058-3113-4).

Stark, J.M., Comstock, J.R., Prinzel, L.J., Burdette, D.W., & Scerbo, M.W. (2001). A Preliminary Examination of Situational Awareness and Pilot Performance in a Synthetic Vision Environment. *Proceedings of the Human Factors & Ergonomics Society 45th Annual Meeting*. Santa Monica, CA: HFES.

Taylor, R.M. (1989). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. *Proceedings of the NATO Advisory Group for Aerospace Research and Development (AGARD) Situational Awareness in Aerospace Operations Symposium* (AGARD-CP-478). Neuilly Sur Seine, France: NATO-AGARD.

Ververs, P.M (1998). *Understanding a Pilot's Tasks*. Savoy, IL: University of Illinois, Aviation Research Laboratory.

Wickens, C.D. (1984). Engineering psychology and human performance. HarperCollins Publishing.

Williams, D., Waller, M., Koelling, J., Burdette, D., Doyle, T., Capron, W., Barry, J., & Gifford, R. (2001). *Concept of Operations for Commercial and Business Aircraft Synthetic Vision Systems*. Hampton, VA: NASA Langley Research Center.

8. Acknowledgements

The research reported herein was supported under NASA contract NAS2-99091 and directed by Dr. David Foyle, manager of the Human Performance Modeling Element of the System-Wide Accident Prevention Project.

The authors thank the following individuals without whom the research reported herein would have been incomplete:

- Pilots Dan Renfroe, Jim Schwartz, Tom Weitzel, and Ken Petschauer for the time and effort they contributed to the ILS task analysis.
- Allen Goodman of NASA Ames for his assistance with organizing the NASA 747-400 simulator scenarios.
- Pilot Rick Shay, a United Airlines 757/767 pilot and SVS test pilot, for his support with the RNAV and SVS cognitive task analyses.
- Ray Comstock and Randy Bailey from NASA Langley for their assistance in collecting information on the Eagle County Regional Airport SVS flight tests.

9. Acronyms and Abbreviations

2D	two-dimensional
3D	three-dimensional
4D	four-dimensional
AGL	(altitude) above ground level
Alt	altitude
A/P	autopilot
APP	approach
ARPT	airport
A/T	auto-throttles
ATC	air traffic control
AvSP	(NASA's) Aviation Safety Program
B727	Boeing 727 aircraft (a medium-sized, tri-jet, 1960s-vintage airliner that carries about 150-180 people depending on configuration)
B757	Boeing 757 aircraft (a medium-sized, twin-engine, 1980s-vintage airliner that carries 200-280 people depending on configuration)
CMD	command
DA	decision altitude
DH	decision height
FAA	Federal Aviation Administration
FAF	final approach fix
FD, F/D	flight director
FL CH	flight level change
FMA	flight mode annunciator
FMS	flight management system
FO	first officer (also known as copilot)

FY	fiscal year (The federal government's fiscal year runs from October 1st to September 30th. So FY02, for example, goes from 10/01/2001 to 9/30/2002.)
GPS	global positioning system
G S	glide slope
GS	ground speed
GSIA	glide slope intercept altitude
HDG	heading
HPM	human performance model (or modeling)
HUD	head up display
IAF	initial approach fix
IAS	indicated airspeed
IFR	instrument flight rules
ILS	instrument landing system
IMC	instrument meteorological conditions
LA	Los Angeles
LAX	Los Angeles International Airport
LNAV	lateral navigation
LOC	localizer
MA	missed approach
MA&D	Micro Analysis and Design (a human factors and human performance modeling small company; see www.maad.com for more information)
MCP	mode control panel
mic	microphone
MSL	(above) mean sea level
NASA	National Aeronautics and Space Administration
NAVAID	navigational aid
ND	navigation display
NM or nm	nautical miles (1 nm = 6076.115 feet)
NY	New York
OM	outer marker (usually also the final approach fix)
PF	pilot flying
PFD	primary flight display
PNF	pilot not flying
RNAV	area navigation (usually using inertial reference systems and global positioning satellite signals)
SA	situation awareness
SBA	Santa Barbara Municipal Airport
SPD	speed
SVS	synthetic vision system
TCAS	traffic alert and collision avoidance system
TFC	traffic
US	United States
VASI	visual approach slope indicator
VERT SPD	vertical speed (climb or descent at a set rate in feet per minute)
VFR	visual flight rules
VMC	visual meteorological conditions
VNAV	vertical navigation
VORTAC	a type of NAVAID

Characterizing Visual Performance During Approach and Landing With and Without a Synthetic Vision Display: A Part Task Study

Allen Goodman
San Jose State University
San Jose, California

Becky L. Hooey
Monterey Technologies, Inc.
Santa Clara, California

David C. Foyle
NASA Ames Research Center
Moffett Field, California

John R. Wilson
San Jose State University
San Jose, California

Abstract

A part-task simulation study of commercial transport approach and landings operations with and without a synthetic vision system (SVS) was conducted in order to generate empirical data and descriptive information for the development of cognitive models of this task environment. Control inputs, eye-tracking, questionnaire, and video recording data were collected from three airline pilots operating over a variety of event and visibility conditions. A description of the simulation configuration, the scenarios tested, and the output data distributed to 5 modeling teams for subsequent analyses is presented. Of primary interest were the eye-tracking data and the observed changes in visual performance associated with the availability of the SVS display. A summary of measures characterizing the distribution, frequency, and duration of the pilots' visual attention during the nominal landing trials (a subset of the scenarios tested) is presented to illustrate the type of analyses that the study data supports. From this summary, regularities in scanning behavior with and without the SVS display are noted, as are localized differences in individual usage strategies.

Introduction

Background

The convergence of several key technologies has accelerated the development and testing of synthetic vision systems (SVS) which provide pilots with an always-available computer-generated perspective view of the outside world ahead of their aircraft. Such a capability affords enhanced spatial and terrain awareness during both darkness and poor visibility. This would especially benefit safety during approach and landing operations when knowing one's position relative to nearby terrain, obstacles, and the runway is always of critical importance. The continued occurrence in both commercial and general aviation of controlled-flight-into terrain (CFIT) accidents and controlled-flight-toward-terrain (CFTT) incidents (Flight Safety Foundation, 1998) attest to this importance.

Though considerable research has been devoted to implementation and design aspects of synthetic vision such as display size, field of view, optimal scene texturing, and appropriate symbology (for summaries, see Purcell, Corker, & Guneratne, 2002; Norman, 2002), less is known about the impact that these systems might have on general flight deck operations. Motivating the present part-task study and corresponding modeling efforts is the recognition that the deployment of new flight deck technologies sometimes alters the manner in which pilots carry out tasks, often in ways that were not fully anticipated by designers (Billings, 1996: Woods & Dekker, 2000). These unanticipated usage patterns can lead to adverse operational consequences.

Clearly, the introduction of a SVS display changes the informational landscape of the flight deck by providing a new and rich source of visual-spatial information. For this reason, an appropriate research issue concerns how pilots will adjust their visual performance to this enlarged landscape. That is, what changes will pilots make in where they look, how long they look, and the sequences in which they look at available informational sources? Gaining a better understanding of how pilots alter their visual performance when flying with a SVS display should help inform the design of effective operational procedures and highlight potential sources of error.

Purpose

The objective of the present study was to generate empirical data and descriptive information regarding pilot performance during approach and landing, with and without the use of a synthetic vision display, for the development of cognitive models of this task environment. (See Foyle, Goodman, and Hooey [2003] in this volume for an overview of the Human Performance Modeling element which this effort supported). The remainder of this paper will describe the details of the part-task simulation study, clarify output data and information made available for model development, and provide an illustrative summary of observed pilot visual performance.

Method

Participants

Three currently flying airline pilots[1] participated in the simulation study. Included in the demographic information collected from the participants was an assessment of prior experience with terrain awareness displays and head-up displays. This was of interest as both display types require pilots to interpret superimposed symbology over terrain information, a task similar to what was required in utilizing the SVS display introduced in this study. Participant demographic information is presented below in Table 1.

Table 1. Demographic information of pilots participating in simulation study.

	Current Crew Position	Current Aircraft Operated	Total Flight Hours	Terrain Awareness Display Hrs.	Head-Up Display Use Hours
Pilot 3	Captain	B-757 / B-767	≈ 16,000	≈ 2000	0
Pilot 4	First Officer	B-747-400	≈ 12,000	0	0
Pilot 5	Captain	B-757 / B-767	≈ 11,000	≈ 1500	≈ 100

Simulator

A PC-based part-task simulator was constructed using the NASA developed PC Plane software package (Palmer, Abbott, & Williams, 1997) which approximated the instruments, controls, and flight characteristics of a Boeing-757. This was linked with an X-IG© 3-D image generation system employing a visual database of Santa Barbara Municipal Airport (SBA) and its surrounding terrain. The simulated flight deck consisted of 6 display components as shown below in Figure 1. Pilot control inputs were made via a joystick with throttle lever, and touchscreen software buttons. A more explicit view of the simulation displays is provided by the photograph shown in Figure 2.

[1] Delineated here as Pilots 3, 4, and 5 which is consistent with other papers in this volume (Pilots 1 and 2 served as simulation check-out pilots and their data is not reported)

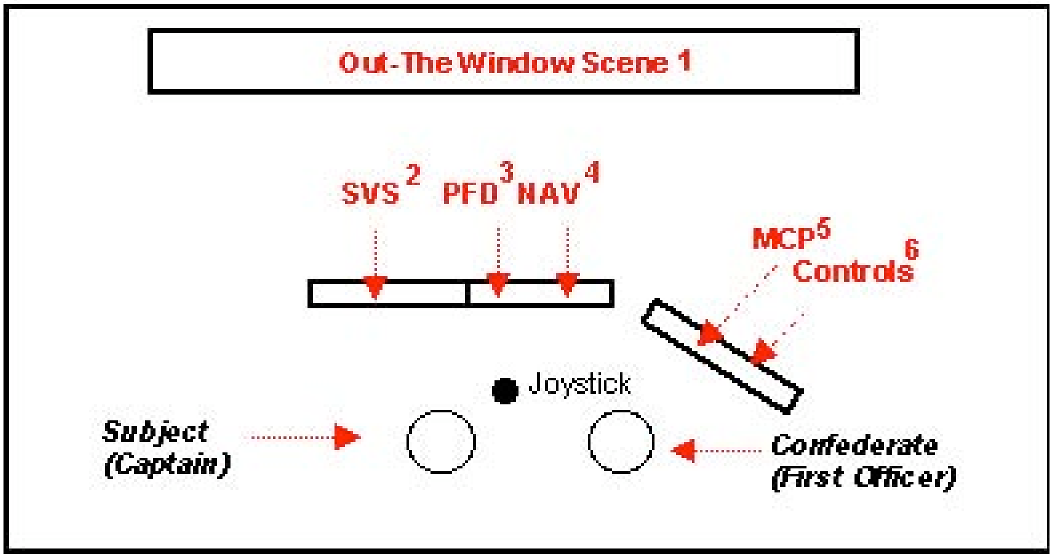


Figure 1. Plan-view diagram of simulation configuration with numbered labels referencing descriptions provided in Displays section of this paper.

Figure 2. Photograph of simulation displays taken from pilot's eye perspective.

Displays

1. Out-the-Window (OTW): The visual out-the-window scene (shown in Figure 3) was presented on a large front projection screen measuring 8 feet horizontal (H) and 6 feet vertical (V) located approximately 8 feet in front of participating pilots. Taking into account the obscuration caused by the front panel of the simulated flight deck, the OTW display provided pilots a 54.6˚ H and 34.9˚ V field of view of the forward external world. On-screen visibility varied between zero visibility and clear visibility according to scenario specifications as detailed in the study design section.

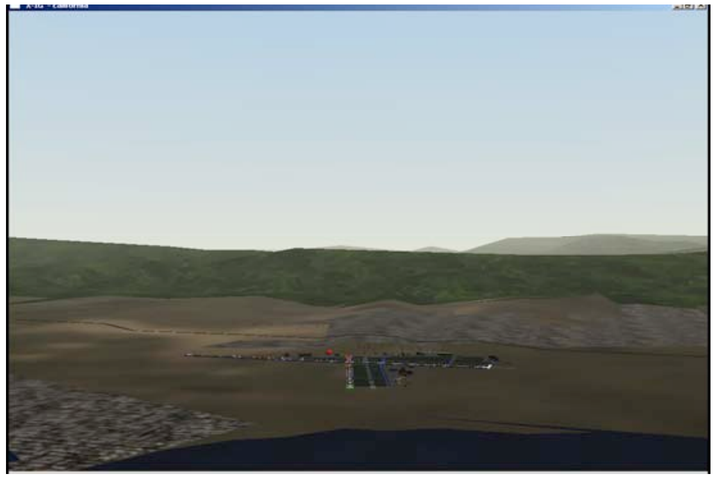


Figure 3. Out-the-Window Scene approaching Runway 33L at Santa Barbara Airport in VMC conditions.

<u>2. Synthetic Vision System (SVS)</u>: The SVS was implemented on a large format head-down display measuring 10 inches H by 7.5 inches V and located 34 inches from pilot eye-point (16.7˚ H, 12.6˚ V). The display presented computer-generated 3-D color imagery of terrain and cultural features overlaid with flight-director symbology, including a flight path predictor (see figure 4). The field of view of the presented imagery was set at 31˚ H and 23˚ V, which provided a somewhat "wide-angle" perspective relative to unity. This fixed field of view was chosen as a good compromise setting, falling between the wide-angle 60˚ horizontal field of view that research had shown (Comstock, Glaab, Prinzel, & Elliot, 2001) was preferred by pilots during early approach phases and the unity field of view desired at landing.

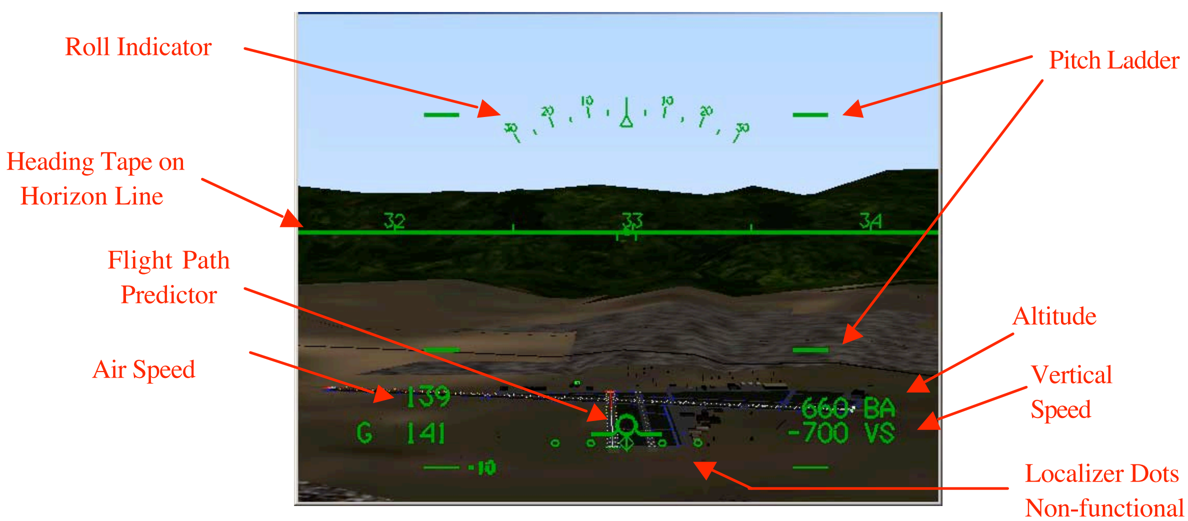


Figure 4. SVS display with identified symbology shown at decision height (DH) on approach to runway 33L at Santa Barbara.

<u>3. Primary Flight Display (PFD)</u>: A conventional primary flight display (see Figure 5) measuring 5.25 inches H and 5.25 inches V was located 34 inches from pilot eye-point (3.1˚ H, 3.1˚ V). The display provided information specifying air speed, attitude, current and

targeted altitude, vertical speed, engine pressure ratios (EPR), distance to next waypoint, and flight mode annunciation

Figure 5. Primary Flight Display (PFD) as seen in simulation.

4. <u>Navigation Display (NAV)</u>: A conventional navigation display (see Figure 6) measuring 5.25 inches H and 5.25 inches V was located 34 inches from pilot eye-point (3.1 deg H, 3.1 deg V). The display provided information specifying ownship track, current heading, latitude/longitude of current position, previous and next fix, distance to next fix, range arcs, and descent crossing arcs.

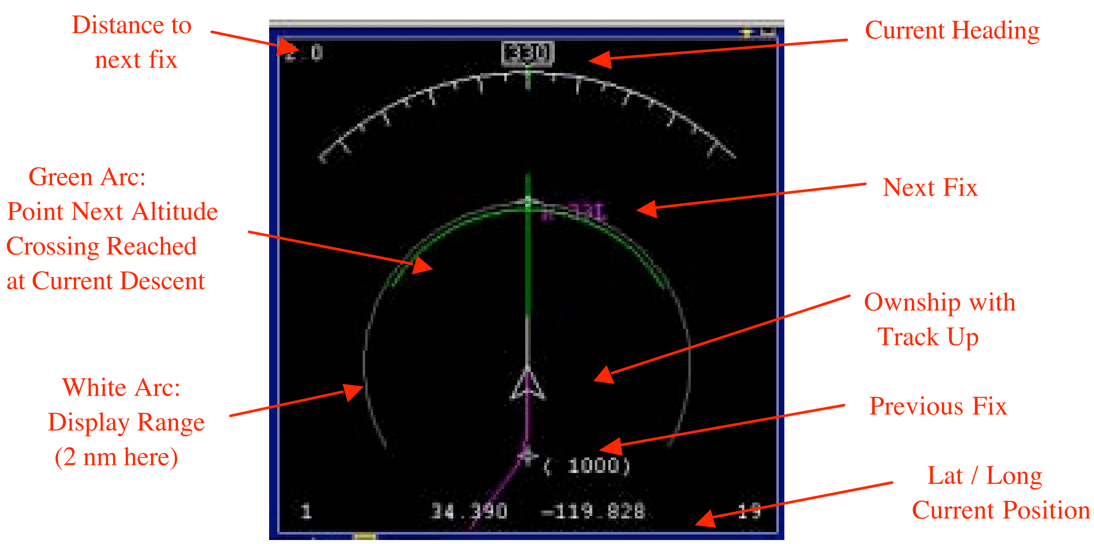


Figure 6. Navigation display (NAV) as seen in simulation.

5. <u>Mode Control Panel (MCP)</u>: A B-757 type mode control panel (see Figure 7) measuring 13.5 inches H and 3.0 inches V was located 39 inches from pilot eye-point (9.2˚ H, 2.0˚ V). The display provided both the status and the means to control autoflight functions through mouse-controlled or touchscreen inputs.

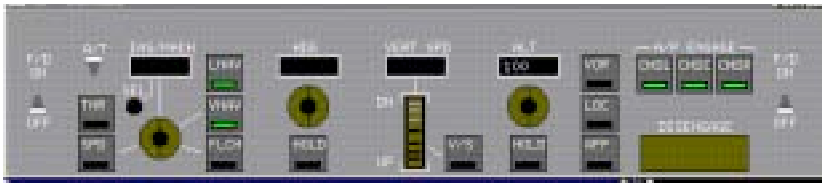


Figure 7. Mode control panel as presented in simulation.

<u>6. Controls:</u>   An unconventional display (see Figure 8) measuring 10.5 inches H and 1 inch V (7.2˚ H, 0.7˚ V) was situated below the MCP. The display provided status information and mouse or touchscreen activation of landing gear, flap setting, speed brake, throttle setting, and NAV map scale.

Figure 8. Controls panel as presented in simulation.

Confederate First Officer (FO)

A flight-qualified member of the experimental study team acted as first officer so as to approximate realistic crew procedures and allocation of duties. These duties included acting on all MCP and control inputs specified by captain, making appropriate call-outs, and handling Air Traffic Control (ATC) communications.

Confederate ATC

A second experimental study team member assumed the role of ATC and provided approach and landing clearances for each trial and, once per testing block, a late reassignment of runway. In no instance did ATC vector aircraft off the programmed route, nor was communications to other aircraft (party line communications) simulated.

General Scenario Description

In all study scenarios pilots performed an Area Navigation (RNAV) daylight approach to Runway 33L at Santa Barbara Airport under calm winds. As this approach does not actually exist, an approach plate was constructed for the simulation based on other published RNAV (GPS) plates (so designated because a Global Positioning System is required) and briefed to pilots. The flight management system (FMS) was preprogrammed by the study team to reflect the approach so that no pilot interactions with the FMS were required during the simulation trials. It should be noted that this type of approach does not require nor make use of ground-based ILS equipment (glideslope and localizer) and represents a trend in future flight operations towards aircraft-based precision guidance.

All trials began 36 nm inbound from the northwest at 10,000 ft and 250 kts awaiting ATC clearance for approach. Pilots were required to fly the cleared approach fully coupled to the autopilot, using the lateral navigation (LNAV) and vertical navigation (VNAV) automated flight modes down to the 650 ft decision height (DH), at which point they took full manual control. Depending on scenario circumstances, pilots either continued the landing (trials terminating at 50 ft) or declared a missed-approach and executed a go-around (trials terminating when ascent reached 3,000 ft). Individual trials lasted approximately 12 min.

Study Design

Independent Variables

Four variables of interest were investigated in the study: *Display Configuration, Visibility, Approach Event, and Phase of Approach.*

Display Configuration

(1) Baseline: This configuration represented current-day operations with the PFD, NAV, MCP, and Controls panel constituting the flight deck displays.
(2) SVS: This configuration included all displays presented in the baseline configuration with the addition of the SVS display.

Visibility

(1) Visual Meteorological Conditions (VMC): The entire trial was conducted in clear day visual meteorological conditions using visual flight rules.
(2) Instrument Meteorological Conditions (IMC): The trial began in instrument meteorological conditions with zero visibility due to dense cloud ceiling down to 800 ft, at which point the aircraft broke-out into clear visibility.

Approach Event

(1) Nominal Landing: The aircraft was cleared first for approach and then landing without incident.
(2) Late Runway Reassignment: The aircraft was cleared first for approach and then landing, but at 1000 ft ATC requested that the pilot conduct a side-step maneuver to the adjacent parallel runway (33R) due to uncleared traffic remaining on runway 33L. (Note: This event was tested in the study during IMC conditions with the SVS display, though current ATC operations would allow such a maneuver only in VMC.)
(3) Missed Approach: Aircraft was cleared first for approach and then landing. In the VMC condition the confederate FO calls out traffic on runway at 600 ft; in the IMC condition dense cloud cover extended to the ground and there was no breakout as anticipated at 800 ft. Both conditions required the captain to declare a missed-approach and execute a go-around.
(4) Terrain Mismatch: The aircraft was cleared first for approach and then landing but instruments (including SVS display) were laterally off-set by 250 ft from the OTW scene. During training, pilots were instructed to declare a missed-approach and execute a go-around if instruments were determined to be unreliable.

Phase of Approach

Phase	Start and End Point	Altitudes	Duration
Phase 1	Start of Trial – Initial Approach Fix	Crossing at 10,000 ft	≈1.0 min
Phase 2	Initial Approach Fix – Final Approach Fix	10,000 ft – 1,800 ft	≈7.5 min
Phase 3	Final Approach Fix – Decision Height	1,800 ft – 650 ft	≈2.5 min
Phase 4	Decision Height – Scenario End	650 ft – 50 ft	≈1.0 min

These four progressive phases of approach differ not only in duration but in external circumstances and required pilot activities. Within Phase 1, pilots are focused on obtaining approach clearance from ATC and setting-up that approach within the auto-flight system. In Phase 2, pilots closely monitor the progress of the approach and configure the aircraft (i.e., set landing gear, adjust flaps and trim). During Phase 3, pilots flying in IMC conditions "break-out" from the cloud ceiling into full visibility and, for all conditions, pilots must visual acquire the runway and confirm proper alignment. By Phase 4, unlimited forward visibility prevails (except in scenarios #5 and #9 as explained below) and pilots must transition to manual control while maintaining proper runway alignment and descent rate.

Test Conditions

The four approach events listed above were flown in three display/visibility configurations: Baseline VMC (current day displays with clear visibility); Baseline IMC (current day displays with dense cloud ceiling to 800 ft); and, SVS IMC (with dense cloud ceiling to 800 ft). This yielded 10 viable test conditions, or scenarios (as shown in Table 1), which were each flown once by the three subject pilots. Collected data from all trials were segmented for analysis purposes into the four progressive phases of flight.

Table 2. Test Conditions (Asterisks [*] denote scenarios from which data is presented in this paper.)

	Display Configuration	Baseline	Baseline	SVS
	Visibility	VMC	IMC	IMC
Approach Event	Nominal Approach (nominal landing)	*Scenario #1**	*Scenario #4**	*Scenario #7**
Approach Event	Late Reassignment (side-step & land)	*Scenario #2*		*Scenario #8*
Approach Event	Missed Approach (go-around)	*Scenario #3*	*Scenario #5*	*Scenario #9*
Approach Event	Terrain Mismatch (go-around)		*Scenario #6*	*Scenario #10*

Procedure

Simulation displays, controls, and procedures were reviewed with participating pilots during an orientation briefing. Thereafter, a training flight was flown in Baseline VMC conditions to familiarize participants with aircraft handling, crew coordination issues, and the Santa Barbara Airport approach. On initiation of each trial, participants were instructed to immediately have the FO arm the autopilots, dial down the altitude, and engage LNAV and VNAV. The participant's task, then, was to monitor and supervise the programmed FMS descent and approach, commanding such actions as flaps, speedbrakes, landing gear, and altitude settings. At DH (650 ft) participants took full manual control (stick and throttle) of the aircraft and either attempted the landing (i.e., descent to 50 ft) or declared a missed approach and executed a go-around.

To minimize pilot adjustment problems in switching back and forth between display conditions, the six Baseline scenarios (without SVS) were randomly divided into two testing blocks of three scenarios each, while the four SVS scenarios were randomly grouped into a third testing block. During data collection, pilots first flew a block of Baseline trials followed by the block of SVS trials and concluded with the remaining Baseline block. Prior to the start of each trial, pilots were told only whether they would be in VMC conditions or IMC conditions (with reported ceiling at 800 ft) and whether the SVS display would be available.

Data Collected

Digital output data were recorded for each trial at 20 Hz across 34 time-referenced variables for aircraft position and orientation, aircraft state, and control inputs.

A helmet-mounted Applied Science Laboratories ASL 5000 eye tracker with eye-head integration was used to collect point of gaze data from participating pilots. Eight sceneplanes were defined in terms of x and y coordinates within a 2-D visual plane encompassing the simulation environment, six of which corresponded to the display regions of interest. For each trial, raw eye fixation data was collected and time-stamped so as to synchronize with other digital data. Data was then processed into a tabular listing of dwell-sequences, durations, and intervals. Summary statistics and sequence probability matrixes were then generated.

For each trial a videotape of the pilot's forward view was recorded from the head-mounted eye tracker. The pilot's point of gaze was shown by crosshairs superimposed over the visual scene. These tapes provided a representation of what the pilot was actually seeing at any given point in the simulation. Additionally, for each trial an ambient audio and video recording was produced that depicted displays and control inputs and verbal communications.

Lastly, after each trial a questionnaire was administered in which subjects rated on a 1 - 7 scale various aspects of their perceived workload and situational awareness across each of the four phases of the just completed flight.

The digital output data, the time synchronized eye-track data, the video recording data, and the questionnaire data collected over the 10 scenarios trials flown in simulation by each of the three participating pilots was distributed, unanalyzed, to five modeling teams. This bundle of empirical data and qualitative information was intended as a primary source of guidance for the construction and validation of cognitive models of pilot performance during approach and landing operations with and without augmented aids. Modeling teams were free to parse and analyze the data as best suited for their particular modeling approach and architectural framework. The specifics of these modeling and simulation efforts are detailed in subsequent papers in this volume.

Selected Results

Of particular interest in this study were the eye-tracking data and what they revealed about changes in visual performance associated with the availability of a SVS display. (The other types of data supplied to the modeling teams are not addressed in this paper). A summary of measures characterizing the distribution, frequency, and duration of the pilots' visual attention during the nominal landing trials is reported below to illustrate the kind of analyses that the eye-track data supports. This summary reflects data collected from scenarios 1, 4, and 7 which differ only in the display/visibility conditions in which participating pilots flew, i.e., Baseline VMC vs. Baseline IMC vs. SVS IMC. All the measures are specified in terms of the six visual regions of interest: OTW, SVS, PFD, NAV, MCP, and Controls. As these six regions constituted the principle sources of flight information within the simulation, fixations recorded outside these regions were ignored. For this summary, the distribution of visual attention was measured as the percentage dwell time in one region relative to the total dwell time in all six display regions. Reported dwell counts indicate the number of "visits" to a particular display region as defined by one or more continuous fixations within a region until visual disengagement. Lastly, reported dwell durations reflect the mean length of time per visit to a display region.

The results for each of the measures are presented graphically (Figures 9a, 9b, 10a and 10b) for each pilot by display condition and phase of approach. The figures are arranged into horizontal panels corresponding to phase of approach and consist of 3 separate graphs, one for each participating pilot. Figures 9a and 9b present the distribution of visual attention of pilots during the nominal landing trials. Here the horizontal axis of each graph shows the display region of interest while the vertical axis represents the percentage of overall dwell time spent in that region. Figures 10a and 10b present the durations and dwell counts. The horizontal axis for these graphs, again, shows the region of interest while the vertical axis represents the mean duration in seconds per dwell. Additionally, a number shown next to each data point marker indicates the dwell count or number of visits made to the display region during that phase of flight.

Based on the observed data, a brief discussion highlighting consistencies and differences in visual performance between approach phases, display conditions, and individual pilots is offered. This discussion is organized in terms of display region.

Out-The-Window Usage

A review of Figures 9a and 9b reveals sizable differences in the percentage of time spent gazing OTW between the three pilots and across the three display/visibility conditions. However, one consistency emerging from the phase of flight data is that regardless of the display/visibility condition, all three pilots devoted an increasing percentage of their visual attention OTW as the aircraft progressed towards landing.

During Phases 1 and 2 in which the aircraft descends to 1800 ft, very little time is committed OTW. Since there is zero visibility for the Baseline IMC and the SVS IMC conditions during these first two phases, it would be expected that pilots only make quick visual checks to see if there is an early break-out from the cloud ceiling. This is suggested in the short OTW dwell durations (mean = .6 sec; see Figure 10a) made by the pilots in this situation. Pilots flying Baseline VMC with clear visibility still relegate no more than 6% of their time OTW through Phase 2.

In Phase 3 the aircraft in Baseline IMC and SVS IMC trials break-out into clear visibility at 800 ft and, for all trials, pilots must visually acquire the runway. Not surprisingly, all three pilots spent the most OTW time in Baseline VMC (10%-18%) in which there was clear visibility throughout the entire phase. However, in comparing Baseline IMC and SVS IMC trials, it is seen that Pilot 3 and 4 each devote 7.5% less time OTW when the SVS display is available. Pilot 5 does not exhibit this characteristic.

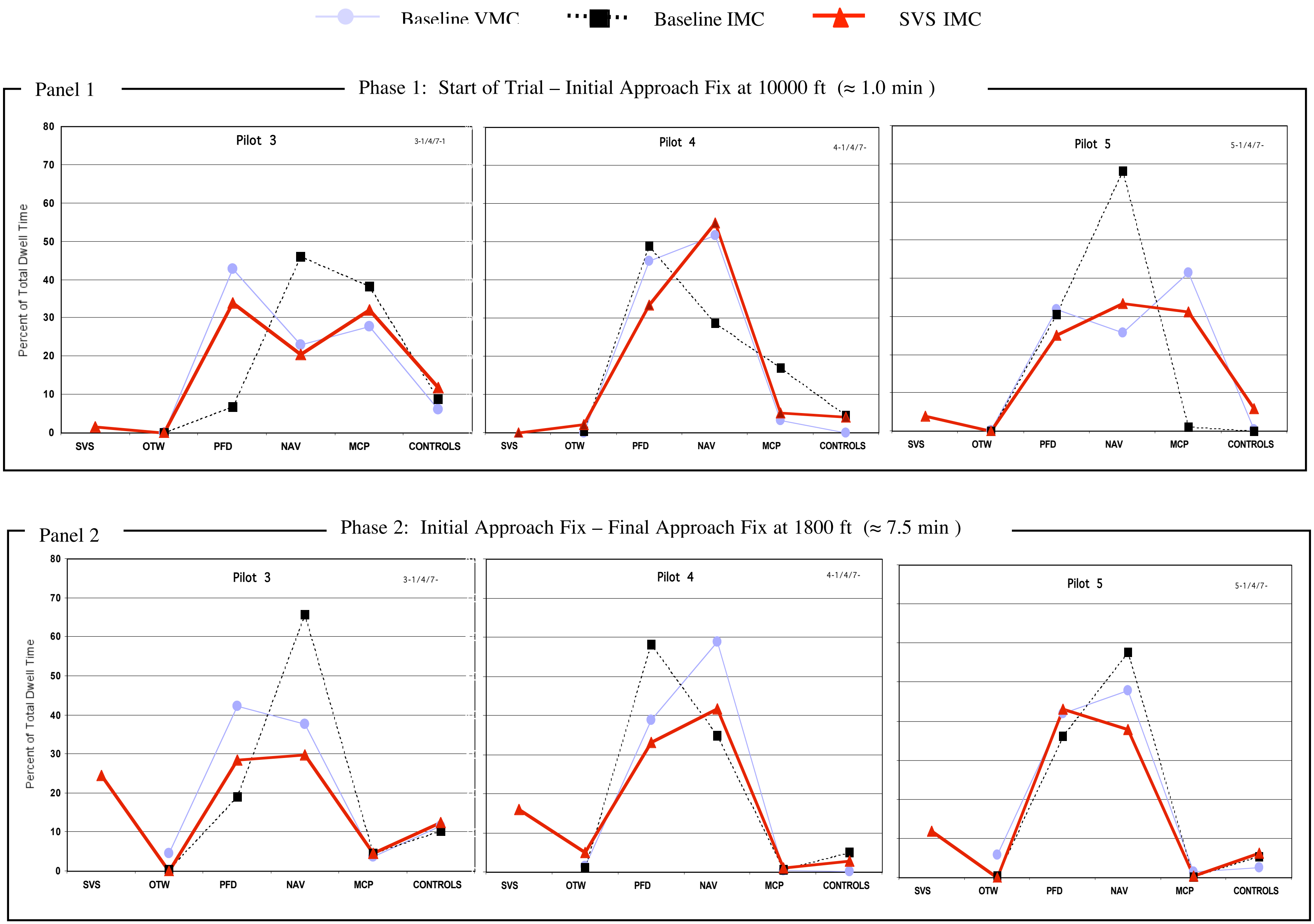


Figure 9a. Pilot dwell time percentages across 6 areas of interest during approach Phases 1 and 2 of nominal landing trials. Three display/visibility configurations are compared: Baseline VMC, Baseline IMC, and SVS IMC.

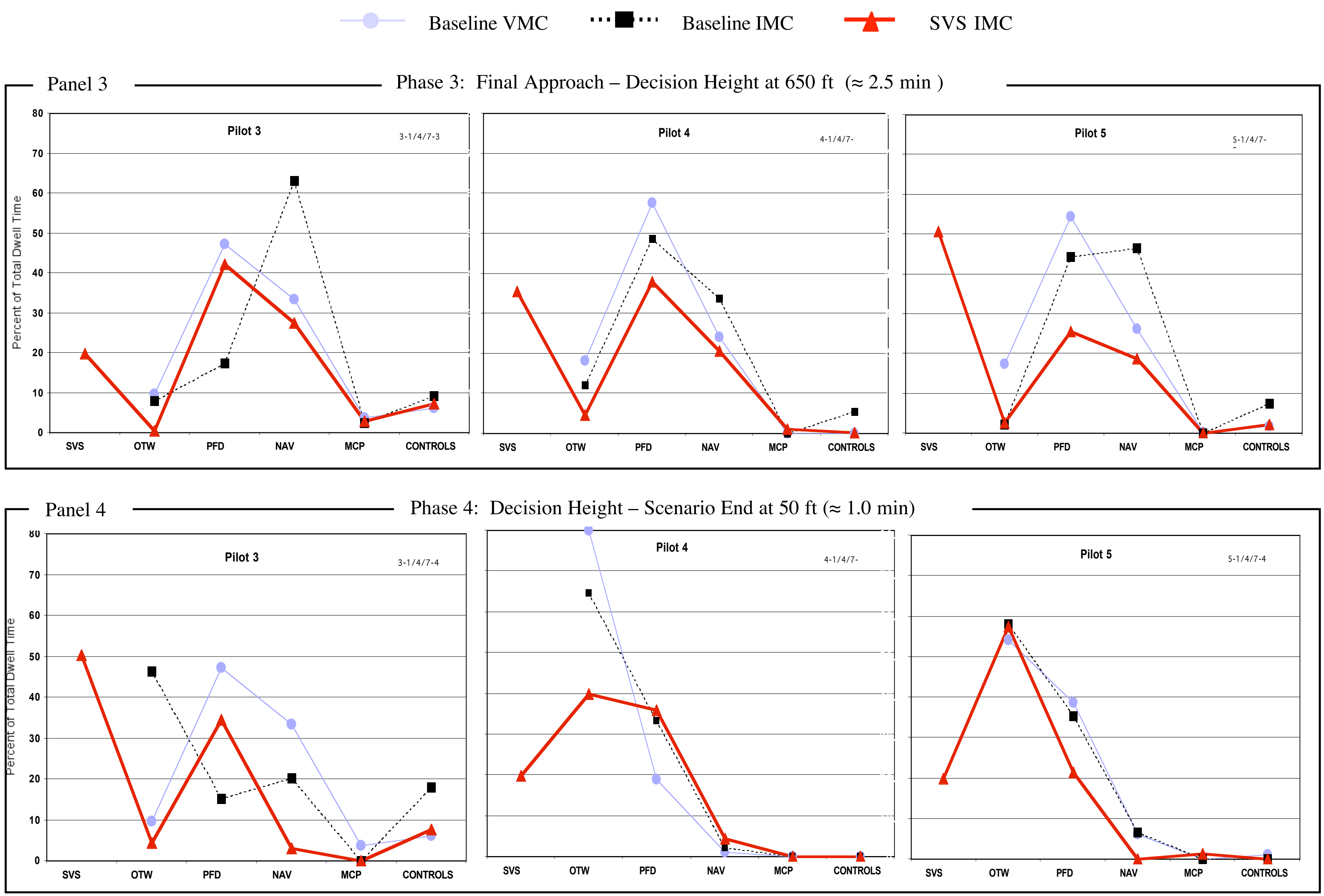


Figure 9b. Pilot dwell time percentages across 6 areas of interest during approach Phases 3 and 4 of nominal landing trials. Three display/visibility configurations are compared: Baseline VMC, Baseline IMC, and SVS IMC.

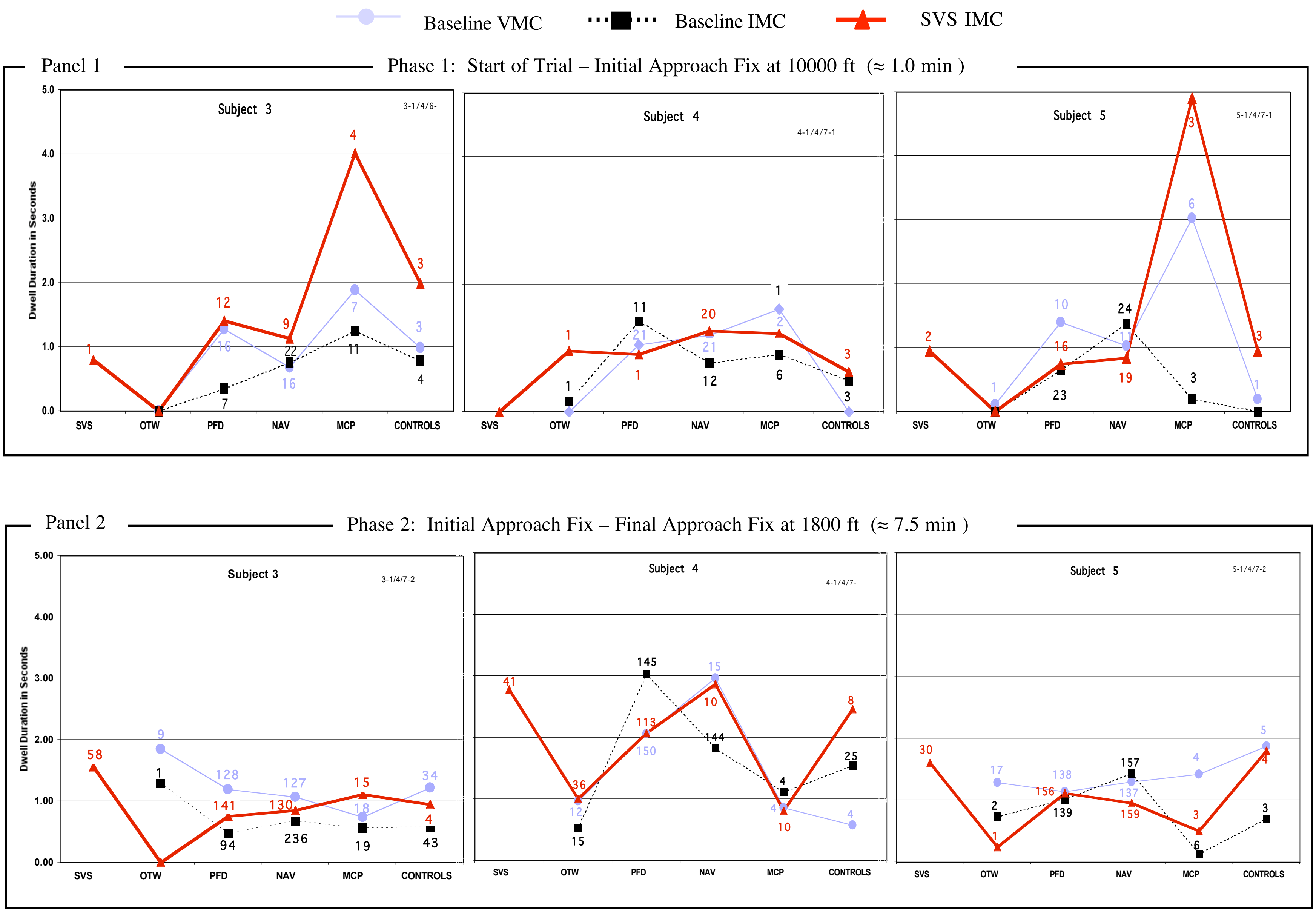


Figure 10a. Pilot mean dwell durations and dwell counts across 6 areas of interest during approach Phases 1 and 2 of nominal landing trials. Three display/visibility configurations are compared: baseline VMC, baseline IMC. And SVS IMC.

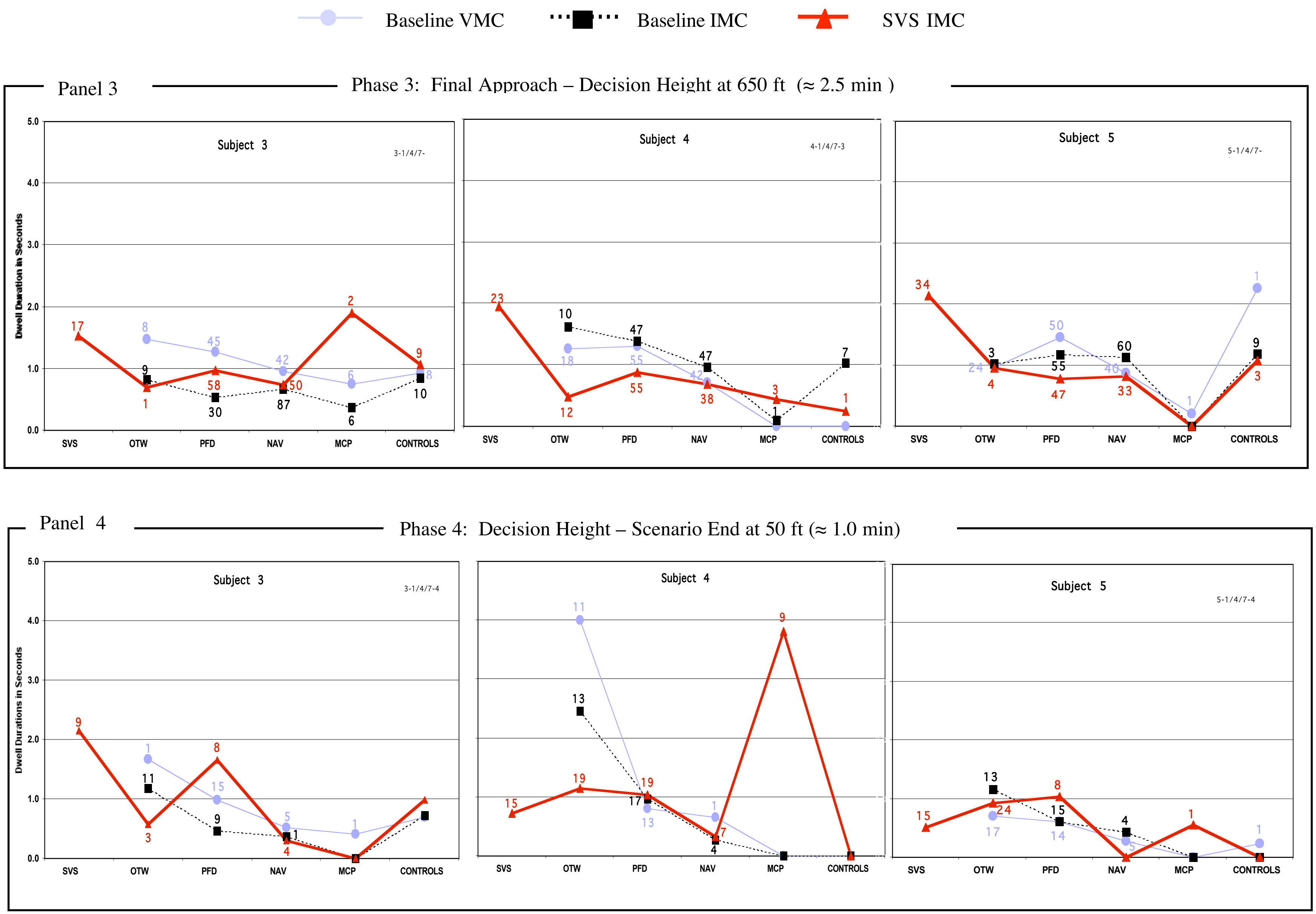


Figure 10b: Subject mean dwell durations and dwell counts across 6 areas of interest during approach Phase 3 and 4 of nominal landing trials. Three display/visibility configurations are compared: baseline VMC, baseline IMC, And SVS IMC.

In Phase 4 with clear visibility across all conditions and an overriding concern for flight path accuracy nearing touchdown, it could be expected that OTW viewing time would be very high and essentially equivalent across conditions. When available, the SVS display would provide no added fidelity over real world viewing, but would offer overlaid flight symbology redundant with the PFD and a flight path predictor useful for assisting landing accuracy. Consequently, it could also be expected that use of the SVS display at this stage would involve quick verification checks of flight symbology which would displace PFD usage and checks of the flight path predictor which, to some extent, might displace OTW viewing.

The data from Phase 4 indicate substantial differences in the visual strategy used by pilots. Pilot 5 best exemplifies the expectations stated above in that OTW viewing occupies the majority of visual attention in this final approach phase and is held nearly constant across conditions (54% - 58%). Pilot 4 devotes even more time OTW in Baseline VMC and IMC conditions, 80% and 65% respectively, but drops down to 40% when operating with the SVS display. For both Pilots 4 and 5, the data show numerous short duration dwells (mean = .6 sec; see Figure 10b) on the SVS display. Pilot 3 demonstrates a different pattern of usage with just 10% OTW time in Baseline VMC and down to just 5% OTW time with SVS. Unlike Pilots 4 and 5, Pilot 3 relies on the PFD and NAV displays and, when available, the SVS display to execute the landing. For this pilot, the usage assumptions regarding the SVS trials are reversed with the OTW view being used as a quick check (mean dwell = .6 sec) of the predominately watched SVS display.

SVS Usage

Pilots 4 and 5 exhibit a similar pattern of SVS utilization over the four phases of approach with little or no usage during Phase 1, increased usage during Phase 2 (14% average dwell time), predominate usage during the break-out and runway acquisition of Phase 3 (43% average dwell time), and, then, reduced usage during the short-final activities of Phase 4 (20% average dwell time). For Pilot 3, SVS usage is approximately constant through Phases 2 and 3 (25% and 20%, respectively), but predominates in Phase 4 (50%).

Comparing differences in dwell time percentages for each of the display regions between SVS and Baseline trials provides a sense of where time spent viewing the SVS display has been diverted. Averaged across the 3 pilots, SVS viewing during Phase 2 reduced NAV display usage 14%, PFD usage 4.5%, and OTW usage .5%. With more time spent viewing SVS in Phase 3, these usage reductions went to NAV display 16%, PFD 10%, and OTW 9%. In Phase 4 the average reductions associated with SVS viewing had shifted to NAV display usage down 9%, PFD usage down 1%, and, interestingly, OTW usage down 18%. It is observed that this last percentage drop in usage, though a mean of all three pilots tested, was the result of the visual performance of Pilots 3 and 4, but not Pilot 5. That two of three pilots selected a strategy of SVS usage during the final phase of approach (actually, the final two phases per discussion above) which substantially reduces OTW viewing is noteworthy.

PFD and NAV Display Usage

The PFD is an important and consistently accessed display with an overall mean percent dwell time across phases, pilots, and conditions of 36%. NAV display viewing figures most prominently during Phases 1 and 2, and when taken together with the PFD, provides the primary source of spatial/navigation awareness to pilots during that period (combined average dwell time of 78%). In Phase 3, when examining just the Baseline conditions, that combined average still remains a robust 81%. But with SVS, combined usage of the PFD and NAV displays drops to 58%. Most of this drop is in the usage of the NAV display which seems to diminish in relative importance with the increased reliance on the SVS display and the nearness of landing.

In Phase 4 the PFD still remains an important source of information and is the second most attended display during this final period. The NAV display, however, is used sparingly (with the exception of Pilot 3 in Baseline conditions) with very short dwell durations (mean = .36 sec).

Concluding Remarks

Pilot control inputs, eye-scan, questionnaire, and video recording data collected in part-task simulation helped guide corresponding modeling efforts by characterizing pilot performance with and without a SVS display during approach and landing operations. Additionally, study scenarios served to specify the task-environment and event conditions to be explored more extensively by the modelers in fast-time simulation. Lastly, performance data from the study provided a means for assessing the validity of resulting model predictions.

Of particular interest in this study were observed changes in visual performance associated with the use of a synthetic vision display. As noted in the selected analyses of the eye-tracking data, these changes included systematic reductions in the dwell times allocated to the NAV, PFD, and OTW displays, mediated by phase of approach. Despite such regularities, there were clear localized differences in SVS usage strategies between the three study pilots. Most significantly, the data shows two of the three pilots spent less time scanning OTW during final approach when the SVS display was available, even though there was unlimited forward visibility. Such usage suggests possible over-reliance on the SVS display.

There is, however, a more fundamental issue which might be derived from these findings. Though participating pilots in this study were made familiar with the functions and features of the SVS display, they were not instructed as to how to utilize the system during test scenarios. In a study of Electronic Moving Map (EMM) usage during taxiing (an advanced, head-down display concept), Graeber and Andre (1999) showed that “uninstructed” participants spent 25% more time viewing the EMM than their “usage instructed” counterparts. This extra head-down, eyes-in time was seen as counter to the potentially performance benefits of EMM and pointed to the importance of procedural training. Similarly, the divergent usage strategies exhibited by pilots in this study highlight the need to develop and train appropriate usage procedures in advance of the deployment of SVS displays.

References

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

Comstock, J. R., Glaab, L. J., Prinzel, L. J., & Elliot, D. M. (2001). Can effective synthetic vision system displays be implemented on limited size display space? *Proceedings of the Eleventh International Symposium on Aviation Psychology*. Columbus, OH: The Ohio State University.

Flight Safety Foundation (1998). Killers in aviation: FSF task force presents facts about approach–and- landing and controlled-flight-into-terrain accidents. *Flight Safety Digest Special Report, Vol. 17, No. 11-12,* 1-256.

Foyle, D. C., Goodman, A., & Hooey, B. L. (2003). An overview of the NASA Aviation Safety Program (AvSP) System-Wide Accident Prevention (SWAP) Human Performance Modeling (HPM) element. In D. C. Foyle, A. Goodman, and B. L. Hooey, (Eds.), *Conference Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays* (NASA Conference Proceedings NASA/CP-2003-212267). Moffet Field, CA: NASA Ames Research Center.

Graeber, David A. & Andre, Anthony D. (1999). Assessing visual attention of pilots while using electronic moving maps for taxiing. *Proceedings of the Tenth International Symposium on Aviation Psychology*, 2, 791-796. Columbus, OH: The Ohio State University.

Norman, R. M., 2002. *Synthetic Vision Systems (SVS) Concept Assessment Report*. NASA Contractor Report NAS1-00106 Task#1002. NASA Langley Research Center, Hampton, VA.

Palmer, M. T., Abbott, T. S. & Williams, D. H. (1997). Development of workstation-based flight management simulation capabilities within NASA Langley's Flight Dynamics and Control Division. *Proceedings of the Ninth International Symposium on Aviation Psychology*, 2, 1363-1368. Columbus, OH: The Ohio State University.

Purcell, K., Corker, K. & Guneratne, E. (2002). *Human factors issues and evaluation for commercial and business aircraft synthetic vision systems*. White Paper prepared for System-Wide Accident Prevention project. Moffet Field, CA: NASA Ames Research Center.

Woods, D. & Dekker, S. (2000). Anticipating the effects of technological change: a new era of dynamics for human factors. *Theoretical Issues in Ergonomic Science, Vol. 1 (3),* 272-282.

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

Michael D. Byrne
Rice University
Houston, Texas

Alex Kirlik
University of Illinois
Urbana-Champaign, Illinois

1. Introduction

This report provides an overview of our ongoing research evaluating the impact of Synthetic Vision System (SVS) technology on pilot performance in commercial aviation in support of the NASA System-Wide Accident Prevention (SWAP) Human Performance Modeling (HPM) element (see Foyle, Goodman, and Hooey, 2003). We first provide a brief discussion in Section 2 of our theoretical and methodological perspective on this problem, and a brief discussion of lessons learned from our research last year modeling pilot performance in the T-NASA Taxi Navigation simulations. We present some of those lessons learned that have provided concrete implications for our current SVS modeling effort.

In the next section of this report, Section 3, we provide an extremely brief description of the NASA/Monterey Technologies, Inc. HPM-SVS Part-Task simulation and experimentation that provided the data set to support our current modeling effort. As indicated there, detailed information on both the simulation and experimentation can be found in Goodman, Hooey, Foyle, and Wilson (2003).

Section 4 of the report provides a description of an extensive set of statistical analyses we performed using the eye movement data collected in the part-task experiments. This section begins with a discussion of what aspects of the overall eye-movement data we selected to focus on for analysis purposes. Next, a set of analyses are presented consisting of, for example, breaking down the data set by phase of flight, approach event (e.g., nominal, missed, terrain mismatch), and of course, the SVS versus non-SVS experimental treatment. As the conclusion of this section demonstrates, these analyses were quite informative in terms of identifying the particular phenomenal that would serve as the focus for our closed-loop, ACT-R modeling.

Section 5 of the report discusses the additional (top-down or theoretical) sources of information guiding our modeling approach, including task analyses, the ACT-R approach, subject matter expert (SME) input, and extant theory of visual attention allocation from both engineering (e.g., Senders, 1964) and psychological (e.g., Wickens, 2002), perspectives.

Section 6 of the report goes on to describe how all the above information led us to focus on particular aspects of modeling pilot performance, and our detailed implementation approach. In particular, that section describes the three central phenomena around which our current efforts have been organized to date: 1) The desire to create a dynamic, closed-loop model of pilot cognition in interaction with the cockpit, aircraft, and environment; 2) The presumption that we are dealing with a relatively knowledgeable and adapted pilot, who is nevertheless presented with novel display technology; and 3) a focus on the allocation of visual attention as crucial to yielding important design- and training-related insights into the impact of SVS technology on cognition and performance. That section concludes with an overview of our detailed, computational implementation as it currently stands, and as we expect it to stand in the near future.

Section 7 describes our findings to date, and the implications of those findings for moving ahead in the near term. The findings are somewhat abstract at this point due to the fact that, although much work has been completed to date, we are only now grappling with the technical issues concerning coupling the ACT-R pilot model with the aircraft model we are using in order to provide a truly dynamic, closed loop account of attention allocation and pilot performance.

The report closes, in Section 8, with a distillation of progress made to date, lessons learned, and future directions. References follow Section 8.

2. Theoretical Perspective and Lessons Learned from Phase 1

2.1 Theoretical Orientation

Aviation incident and accident investigations often find both cognitive and environmental sources of human error. Environmental sources include factors such as flawed interface design, confusing automation, and unexpected weather conditions. Improved environmental design, such as the use of the Synthetic Vision Systems (SVS) that are the subject of our current research, often provide important leverage for reducing error and improving human performance. On the other hand, cognitive sources underlying the effectiveness and efficiency of performance include factors such as situation awareness, procedural compliance or non-compliance, and crew coordination. Many if not most significant incidents and accidents result from some combination of both cognitive and environmental factors. In fact, in a highly proceduralized domain such as aviation, with pilots who are highly trained and well-motivated, accidents rarely result from either environmental or cognitive causes alone. Training and experience are often sufficient to overcome even the most confusing interface designs, and the environment is often sufficiently redundant, reversible, and forgiving so that the vast majority of cognitive slips and mistakes have no serious consequences. Most highly consequential incidents and accidents result only when both environmental and cognitive factors collectively conspire to produce disaster.

Introducing new technology is a common approach to trying to reduce either the frequency, severity, or consequences of less-than-perfect pilot performance. Human performance modeling associated with evaluating the impact of technological interventions therefore requires giving consideration to both cognitive and environmental issues. This report describes the progress made to date on a research project in which dynamic, closed loop cognitive-environmental modeling, or more specifically pilot-vehicle-airport modeling, is currently being performed in order to shed light on both the positive and potential negative effects on the introductions of SVS technology in the commercial airline cockpit. Our current modeling consists of integrating a pilot model developed within the ACT-R cognitive architecture (Anderson, Bothell, Byrne, & Lebiere, 2002) with a commercial, off-the-shelf (COTS; Bowers and Jentsh, 2001) model of aircraft dynamics and the Santa Barbara airspace and airport which served as the basis for part-task experimentation. The overall objective of the NASA program in which we are participating is to develop computational human performance models with the predictive ability to aid designers and analysts in identifying likely vulnerabilities in human-machine performance in aviation.

Our current SVS modeling is actually the second stage in a longer term effort to meet this goal. Our research in the previous year focused on modeling to understand the causes of taxiing errors in a NASA simulation involving taxi-to-gate scenarios at a simulation of Chicago O'Hare (ORD) airport. To set the stage for the rest of this report, we briefly discuss the lessons learned from that effort which motivated our approach to the current problem. These lessons helped us get a bit of a head start regarding the selection of an initial modeling architecture for SVS modeling: a closed-loop, dynamically interacting dyad comprised of an ACT-R model of cognition and a commercially available aircraft-airport simulation package.

2.2 Lessons Learned From Phase 1

As we learned in our taxi modeling research, it is a nontrivial matter to apply scientific models of cognition, developed and validated primarily with psychological laboratory data, to applied contexts such as human-machine performance in aviation. Specific challenges include the following.

2.2.1 Communication Between Cognition and the Outside World

Experimental tasks are typically carefully designed in such a way that the inputs to, and outputs of, the cognitive system are readily identifiable. This is largely done by making the perceptual and motor demands associated with cognitive experimentation relatively trivial. Unfortunately, the perceptual and motor demands associated with aviation cognition can be extensive. This problem surfaced in Phase 1 research in the difficulties associated with coupling the ACT-R/PM model with the visual scene database, and also to some extent, with the aircraft model. The latter was not a severe problem because motor outputs could be considered to be relatively discrete in this instance, consisting of distinct settings of the throttle and brake. The SVS scenarios are similar in this regard although, as will be seen below, we have dedicated a good bit of effort to couple the perceptual mechanisms of ACT-R with both the cockpit and external scene provided by the flight simulator with which ACT-R is intended to interact.

2.2.2 Modeling Environmental Objects and Dynamics

As Phase 1 research clearly demonstrated, achieving a reasonable model of pilot cognition in dynamic, interactive contexts depends heavily on the availability of reasonable models of the visual, physical, and controlled environment, as well as its dynamics. The dynamics of human cognition and behavior is interleaved with, and occurs in concert with, the dynamics of environmental entities that also participate in the functioning of the integrated human-environment system. In our Phase 1 final report, we noted that cognitive modeling software packages can make better and more explicit provisions for representing objects and dynamics in the external environment so facilitate the task of modeling interactive behavior in contexts more complex than the desktop computer. As will be seen below, this issue has again resurfaced as a non-trivial issue in our current SVS modeling, if not theoretically, at least from a technical and data-communication perspective.

2.2.3 Timing Issues

On the basis of our Phase 1 research we concluded that modeling systems for running human-environment system models should provide separate clocks and processing resources for simulating cognitive and environmental dynamics. We noted that neither is subservient to, nor should be subsumed, by the other, nor should either model have to wait while the other is updating. We additionally noted that the current solution to this problem, in which processing resources are passed back and forth between cognitive and environmental components of the model, will be found to be increasingly unwieldy as more dynamic contexts are modeled. On the basis of this finding from our Phase 1 research, we made an early and explicit decision to implement our ACT-R pilot model and our aircraft/airport model on separate computers, connected by networked resources for 2-way data communication. With this approach, the processing demands associated with mimicking the dynamics of the two components of the total interactive systems do not compete. However, we have learned that neither of the two pieces of software we are using for modeling environments make explicit provisions for such inter-computer communication. As a result, we are currently at the stage of engineering our own solution to this problem.

3. Part-Task Simulation and Experimental Design

The focus of this research is on the evaluation of a new technology for the commercial airline cockpit (Foyle, Goodman, and Hooey, 2003). One of the factors that has long limited aviation is visibility; poor visibility conditions can substantially change the task of piloting an aircraft. However, with extensive and accurate computer-based geographic information systems, it is possible to generate the view of known terrain as long as the location of the observer is known. Modern GPS systems make it possible to know the location of an airplane with high accuracy. Thus, the combination of the two systems makes it possible to render on a computer display the terrain that may not be visible due to adverse environmental conditions (e.g. fog, rain). This is the basis for NASA's Synthetic Vision System or SVS (for detailed information on the SVS design used in the current study, see Goodman, Hooey, Foyle, and Wilson, 2003).

Thus, a Synthetic Vision System (SVS) is essentially a computer generated display designed to provide the pilot with information that augments the out-the-window view, to better enable the pilot to fly safely, at low levels, through traffic, around terrain, and in low visibility conditions. Experiments conducted at NASA Ames Research Center by NASA and Monterey Technologies Inc. were performed to investigate the potential positive and negative effects of augmenting a cockpit with a prototype SVS display (Goodman, Hooey, Foyle, and Wilson (2003).

The purpose of these experiments was to collect data characterizing pilot performance and eye movement behavior during the approach and landing phase of flight using with both conventional and augmented displays under both Instrument Meteorological Conditions (IMC) and Visual Meteorological Conditions (VMC) conditions. The experimental plan, due to the cost and time required for studies of this complexity, focused on a limited number of pilots operating across a variety of conditions and treatments, as described in Goodman, Hooey, Foyle, and Wilson (2003).

4. Eye Tracker Data Collection, Interpretation, and Analysis

The first issue to be considered in this modeling effort is what to model. That is, what variance is there to be explained? In these scenarios and with the limited number of subjects available, there is not a rich corpus of speech or control manipulations, let alone errors, to model. Indeed, even if there were more errors, it is not clear how representative the small sample would be. The sample is also limited in terms of airports, weather conditions, scenarios, etc.

However, there is one extremely rich source of data: the eye movement record. Not only is there a great deal of data with which to work, there is substantial variance to be explained. Furthermore, we believe this eye movement record is the most generalizable component of the data. What we want to know is how the SVS, a visual display, affects pilots' visual behavior. For instance, if some higher-level metric of task performance (such as number of errors made) is not sensitive enough to show an effect of the SVS, this may be because the SVS is ignored by the pilots or because the difference it makes is compensated for by other factors. We can distinguish those cases via the eye-tracking data. Thus, we focused our empirical analysis on the eye data.

4.1 Collection and Coding Issues

Raw eye tracker data was provided by NASA at 20 Hz without any filtering or smoothing of the data. The data is noisy, which may be due to several sources such as blinking or a temporary loss of eye tracker calibration. The raw data are very detailed, containing raw X and Y coordinates, pupil dimensions, and other information. We looked primarily at one variable, region of interest (ROI), which was divided into eight sceneplanes, as follows (see also schematic in Figure 1):

Sceneplane 0 = Undefined or invalid data. Occurs when the eye cursor is centered on an area that is not defined as sceneplane 1 to 7 – i.e. the first officer, joy stick etc - or if the data is invalid (i.e. subject blinks).
Sceneplane 1 = Out-the-Window (OTW) View
Sceneplane 2 = SVS Display

Sceneplane 3 = Primary Flight Display
Sceneplane 4 = Nav Display
Sceneplane 5 = Mode Control Panel
Sceneplane 6 = Controls (Flaps, gears, speedbrakes, map scale)
Sceneplane 7 = Overlapping Area. The cockpit displays sit directly in front of the lower portion of the OTW view. Depending on the viewing angle of the subject (which varied slightly by subject, and over the day of trials), the eye tracker could not always determine whether the subject was looking at the black masking area around the displays, or the OTW view behind the masking. In these cases, the sceneplane was recorded as “7”.

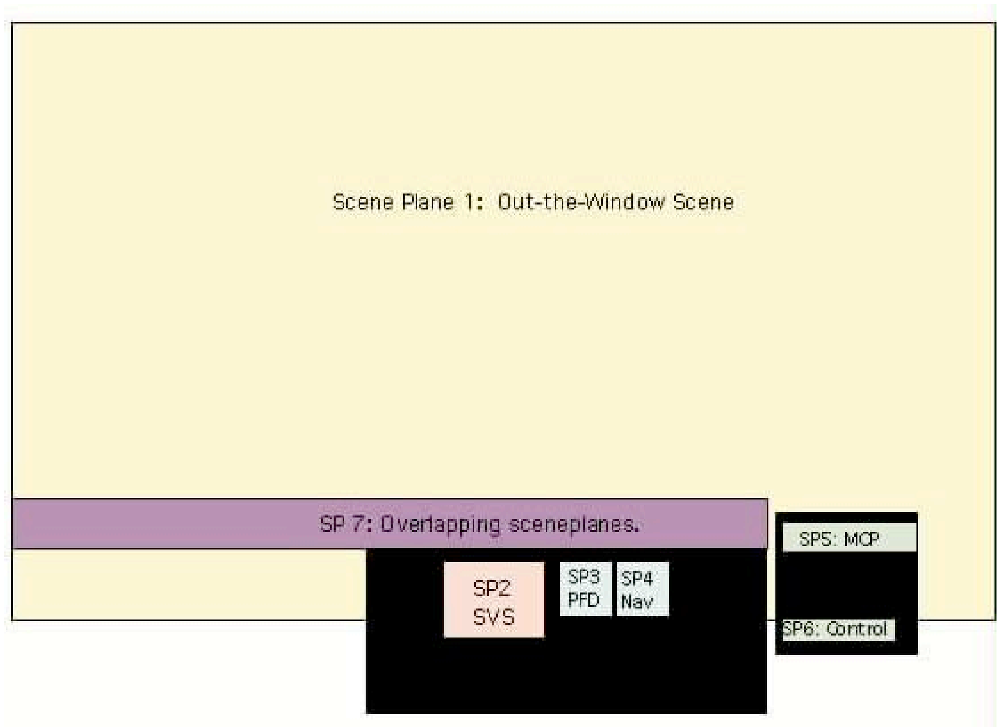


Figure 1. Overview of Sceneplane Layout.

In addition to the raw tracker data, NASA provided videotapes from two cameras:

Eye Tracker Camera. For each trial a videotape of the pilot's forward view was recorded from the head-mounted eye tracker. The pilot's point of gaze is shown by crosshairs superimposed over the visual scene. These tapes provide a fair representation of what the pilot was actually seeing at any given point in the trial.

Room View Camera. Additionally, for each trial an ambient audio and video recording was produced that depicts displays and control inputs and verbal communications. Three audio channels were recorded as follows: left channel was the Captain (subject), right channel was the FO (experimenter), and center channel was ATC (experimenter). It should be noted that the camera was mounted high and behind the pilot and that the visual perspective in the tapes is not that of the pilot.

These additional sources of data were useful in understanding what the pilots did and when they did it for the purposes of validating the task analysis, but were not quantitatively analyzed.

4.2 Eye Movement Analysis and Results

Fixations are distinguished from saccades (rapid voluntary eye movements used to move from one fixation to another) and very small involuntary eye movements of several types that occur during fixation. A "dwell" is defined as the time period during which a fixation or series of continuous fixations remain within a sceneplane or ROI.

Selected Data Analysis

To focus on SVS versus non-SVS cases in similar conditions, our initial data analysis focused solely on the conditions listed in Table 1.

Table 1. Selected conditions for data analysis

Display Configuration		Baseline	SVS
Visibility		IMC	IMC
Approach Event	Nominal Approach (nominal landing)	*Scenario #4*	*Scenario #7*
	Missed Approach (go-around)	*Scenario #5*	*Scenario #9*
	Terrain Mismatch (go-around)	*Scenario #6*	*Scenario #10*

Because we are primarily concerned with the allocation of visual attention, the primary variables of interest are those that index the amount of attention given to each sceneplane. Allocations can be counted in two ways, by count (number) of fixations or by total duration of dwell. Of course, if all fixations have the same duration, the relative proportions spent on each sceneplane will be identical, but this is an empirical question.

Additionally, there are a potentially large number of ways to segment the data. In this presentation we will limit our analysis to the following: we consider only four categories, namely, pilots' baseline eye movements, pilots' SVS versus baseline eye movements, pilots' eye movement by different flight phases, and pilots' eye movements by different approach scenarios.

4.2.1 Baseline Attention Allocation

The first thing to be established is the baseline allocation of attention. If the SVS is to serve as a proxy for looking out the window, it is useful to know how often pilots look out the window in the first place. In addition, since the SVS contains information that can be found on other displays (e.g., altitude, which is also on the PFD), it is important to know how often those other displays are accessed.

Figure 2 shows that overall pilots spent 40.2% of their total fixation dwell time on the PFD, and about 40% fixation duration on NAV. They only spent about 3% looking out the window, about 4.2% at the MCP, and about 6.5% at the Controls. This result suggests that the PFD and NAV displays are the major targets of attention, as they account for about 80% of the fixation dwell time.

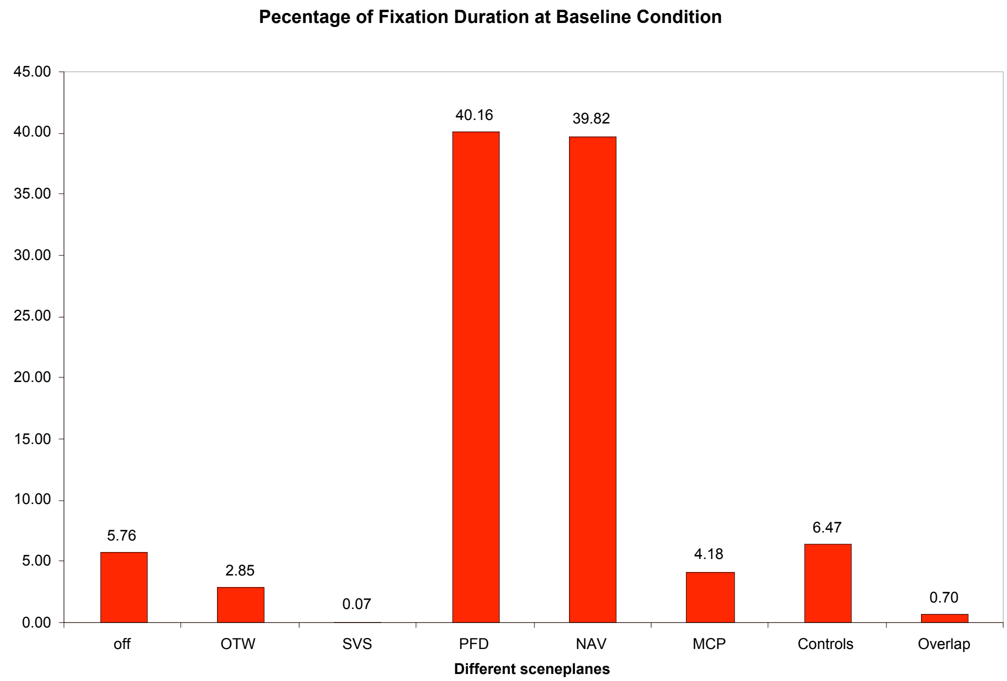


Figure 2. Percentage of fixation dwell time in the baseline condition.

The fixation count measure shows a similar distribution, as depicted in Figure 3. Pilots spent 38.3% of their fixations on the PFD, and another 39.6% of their fixations on NAV, and relatively small amounts of time looking at other displays, most notably less than 3% out the window.

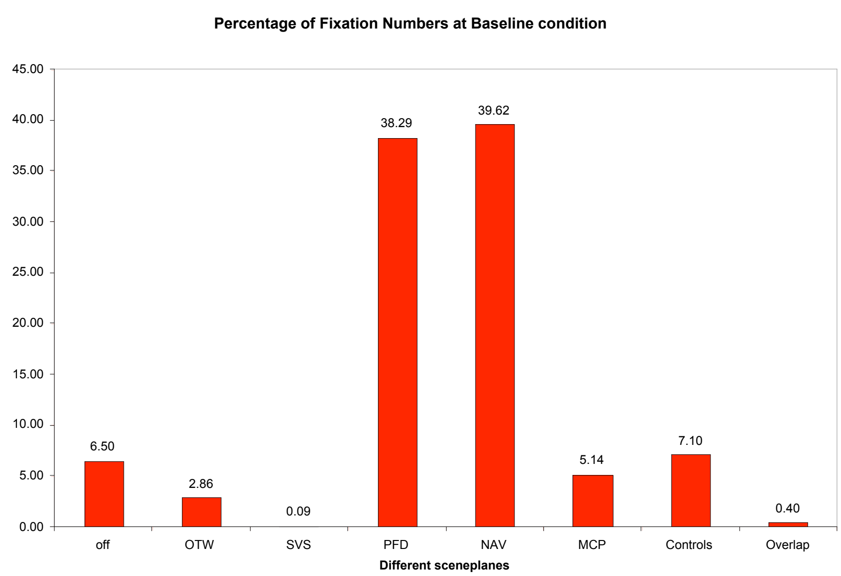


Figure 3. Percentage of fixation count in the baseline condition.

4.2.2 Pilots Eye Movements at SVS versus Baseline Conditions

The obvious next question is how the SVS affects this distribution. Figure 4 shows that the distributions of pilots' fixation dwell time on different sceneplanes are similar for both the SVS and the baseline (No SVS) condition. In the SVS condition, pilots also spent most of their fixation time on the PFD (35.7%) and NAV (29.6%). Pilots also spent 2.6% fixation time looking out the window in the SVS condition. However, one notable feature in the SVS condition is that pilots displayed a large number of fixations (20.2%) on the SVS. This clearly shows that pilots did look at the SVS display quite frequently.

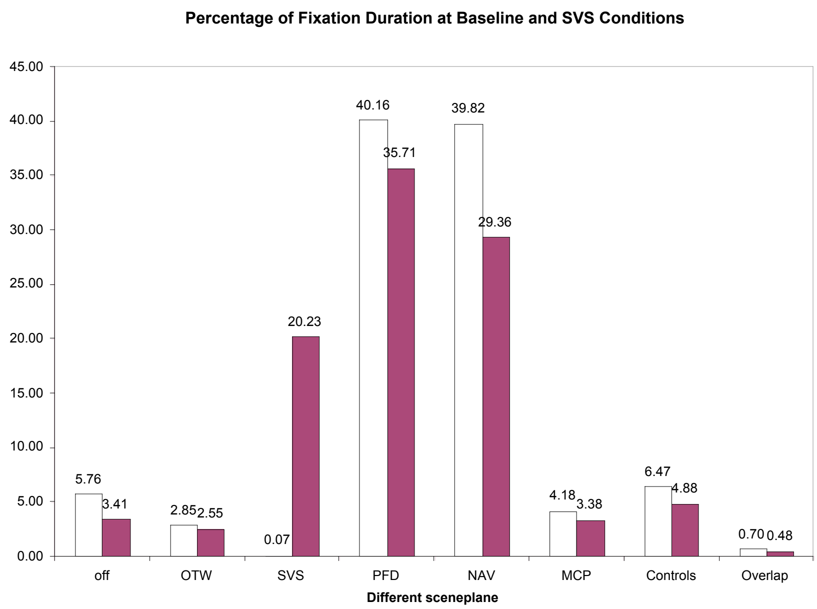


Figure 4. Percentage of total fixation dwell time on each sceneplane in the baseline (light bars) and the SVS (dark bars) conditions.

Again, the distributions for fixation count follow those for time. Figure 5 shows the graph for fixation count for SVS and baseline conditions, and yields about the same results.

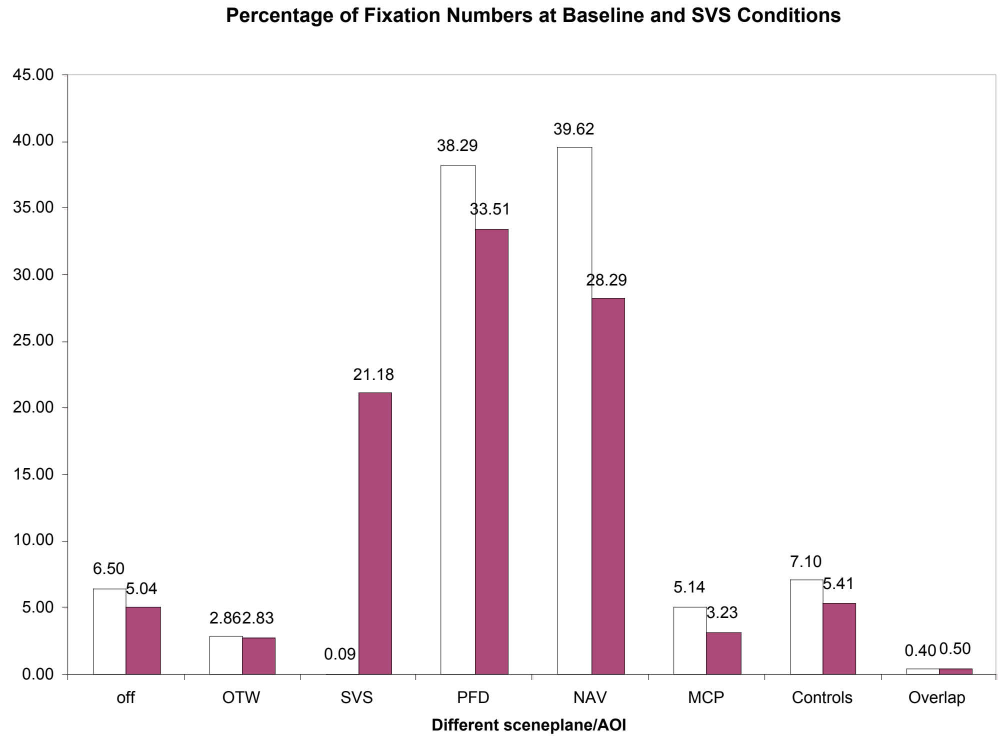


Figure 5. The percentage of fixation count on each sceneplane in the baseline (light bars) and the SVS (dark bars) conditions.

Clearly, pilots spend a fair proportion of their gaze on the SVS. An interesting question we wanted to address is from where they "stole" these gazes. Namely, they must reduce some portion of fixations associated with other sceneplanes. A natural suspect for the location from which fixations would be stolen was the OTW display, due to the inherent redundancy between the perceptual information gained from the SVS and OTW displays.

Surprisingly, however, the fixation dwell time and frequencies on different sceneplanes under SVS versus baseline (No SVS) conditions, as shown in the above two figures, shows no obvious difference in the amount of gaze directed out the window. Instead, it appears the SVS is associated with a reduction in the amount of gaze directed at the PFD and NAV displays. Thus, the SVS display was drawing attention away from other sources of information from within the cockpit, it was not acting as a "substitute" for the information provided by the OTW display.

We find this result both counterintuitive and quite interesting. The underlying rationale behind the SVS display is that it will act as a substitute for the OTW display when the information obtainable from the latter is degraded. Instead, these data seem to indicate that it did not act as a substitute source of environmental information (at least for these pilots). And as a possibly unintended result of the presence of the SVS, less attention was paid to other displays.

4.2.3 Analysis by Phase of Flight

All the above analyses combined all the flight phases together. During different flight phases, however, pilots may have different needs for different information. We therefore decided to break down the analysis of data by flight phase. The fight phases were defined as follows:

Phase 1. Start to Initial Approach Fix (IAF)
Phase 2. Initial Approach Fix (IAF) to Final Approach Fix (FAF)
Phase 3. Final Approach Fix (FAF) to Decision Height (DH)
Phase 4. Decision Height (DH) to End

Because the percentage of fixation dwell time and fixation counts on different sceneplanes provides similar information, to simplify the analysis we only dealt with the percentage of fixation dwell time on different sceneplanes in all subsequent analysis. Again, the natural starting point is the baseline condition, which is presented in Figure 6.. Note how the use of the PFD increases as the flight moves on, and the sharp drop in the use of NAV and sharp increase in OTW gazes in phase 4.

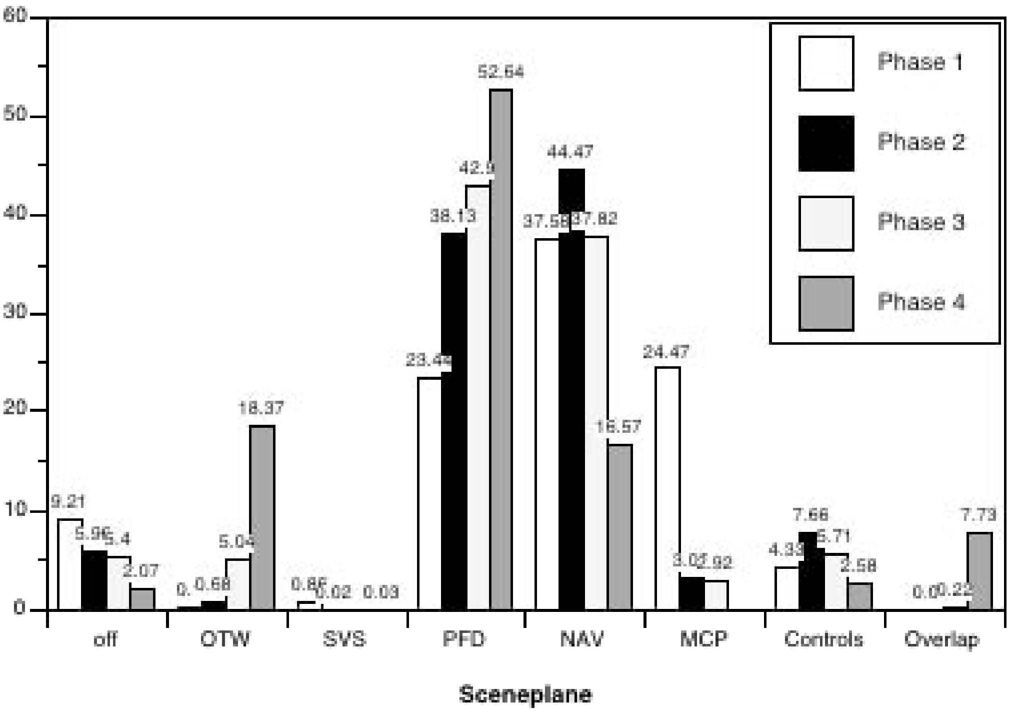


Figure 6. Percentage of fixation dwell time in the baseline condition by different flight phases.

Because pilots so rarely look out the window in phase 1, one might expect little use of the SVS in this phase. In fact, pilots do look a little at the SVS in this phase but overall the allocation of gaze is not substantially different in baseline vs. SVS in this condition, as shown in Figure 7.

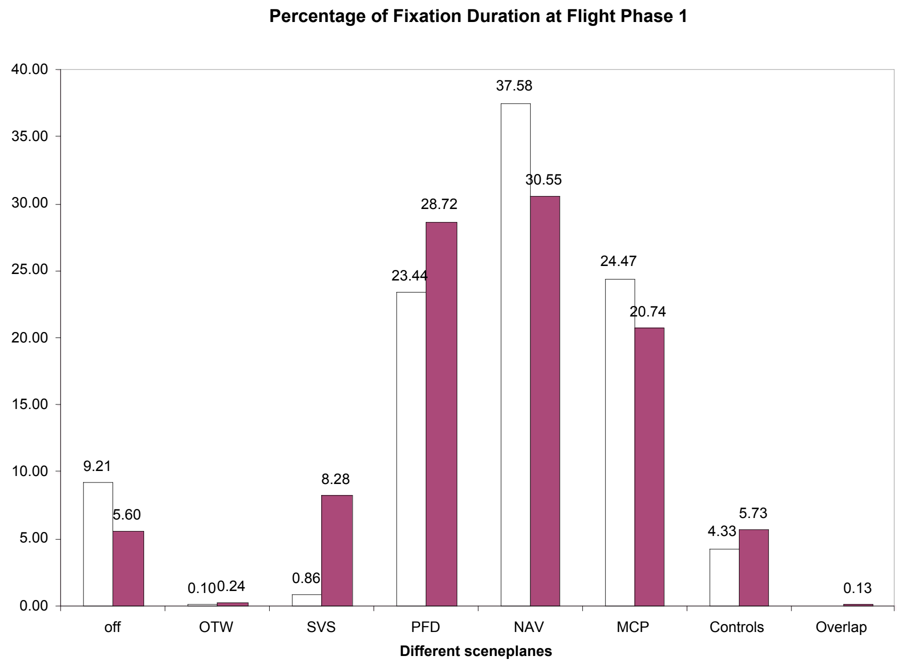


Figure 7. Percentage of fixation dwell time in the baseline (light bars) and the SVS (dark bars) conditions for flight phase 1.

.

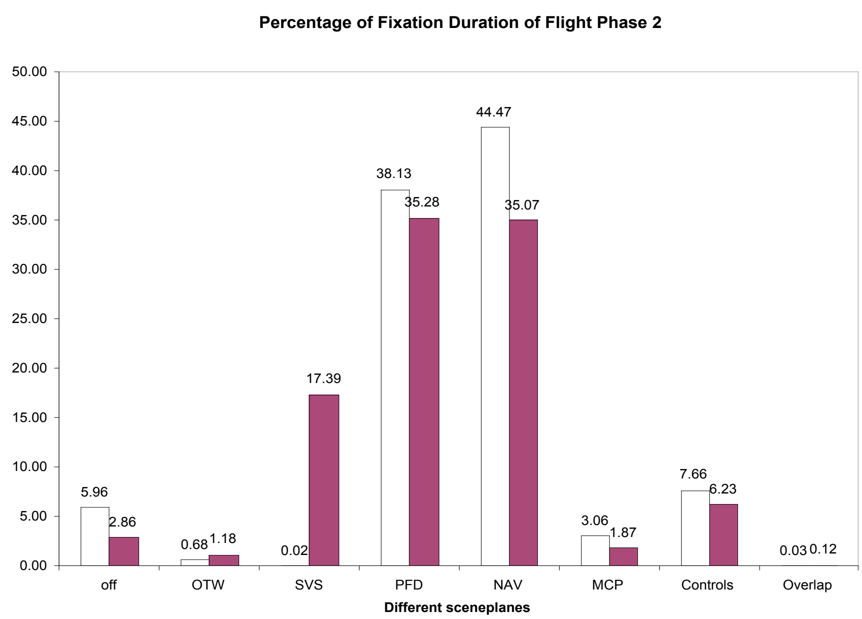


Figure 8. Percentage of fixation dwell time in the baseline (light bars) and the SVS (dark bars) conditions for flight phase 2.

Phase 2 shows an increase in SVS use over phase 1, with most of the gaze time being "stolen" from PFD and NAV (see Figure 8). Note that in this phase of flight, there is virtually no use of OTW in either baseline or SVS conditions, but the SVS is still used when present. We believe this is due primarily to the symbology overlaid on the SVS. This trend continues through phase 3, depicted in Figure 9. Note the substantial use of SVS and reductions in use of PFD and NAV

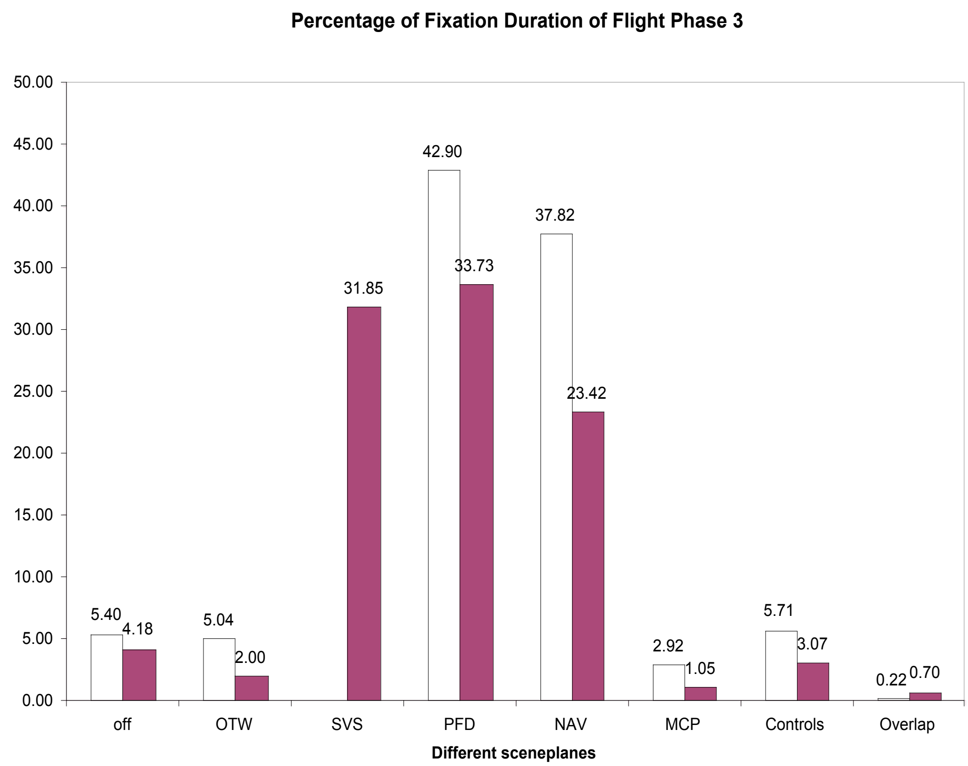


Figure 9. Percentage of fixation dwell time in the baseline (light bars) and the SVS (dark bars) conditions for flight phase 2.

Flight phase 4 shows the largest reduction of OTW looking as a result of the SVS (Figure 10), but this reduction in OTW gaze only accounts for about a third of the total SVS time, which appears to come primarily from the PFD. Overall, SVS gaze accounts for a fairly substantial proportion of total gaze time in phases 3 and 4, but this is not by simple reduction of OTW gaze. Instead, pilots seem to borrow gaze from the PFD.

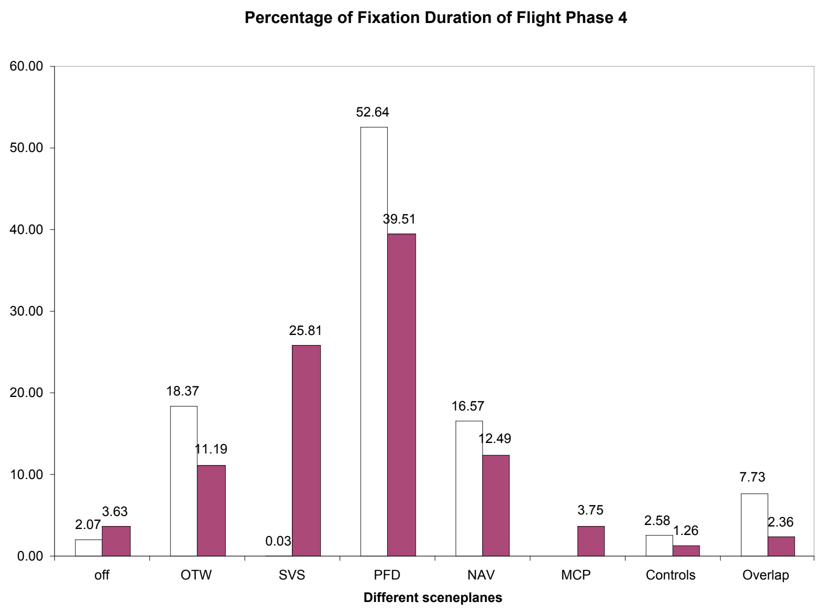


Figure 10. Percentage of fixation duration at the baseline and the SVS condition of flight phase4.

4.2.4 Late Phases of Flight Broken Down by Approach Events

Since the SVS versus non-SVS differences showed up most prominently in the final phases of flight, the next step in the data analysis focused on phase 3 and phase 4, further breaking down the data by different approach scenarios, namely, Nominal Landing, Missed Approach and Terrain Mismatch. This is depicted in Figures 11-13.

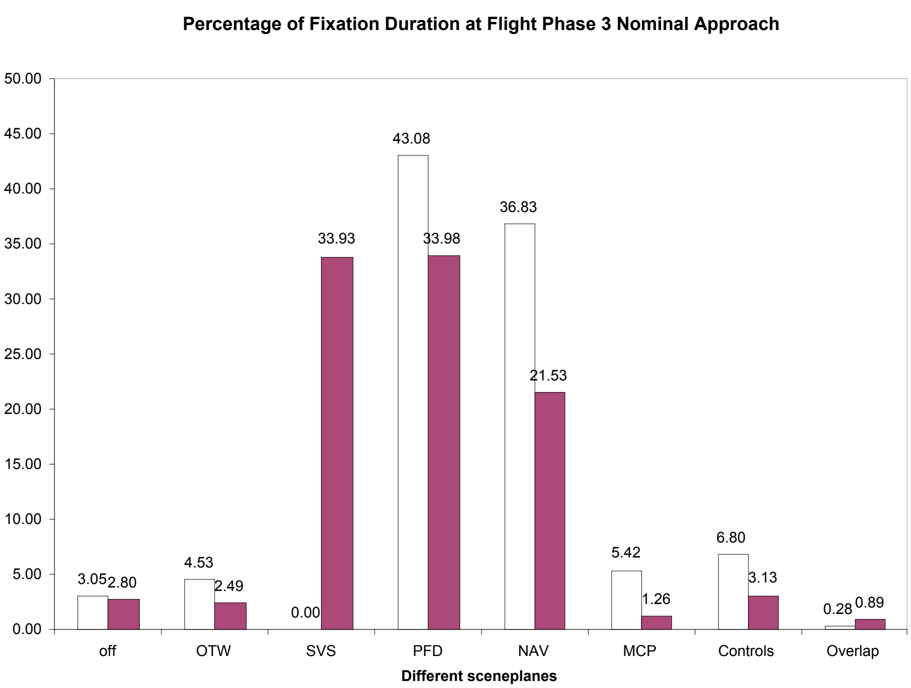


Figure 11. Percentage of fixation dwell time in baseline (light bars) and SVS (dark bars) conditions during flight phase 3 for the Nominal Approach scenarios.

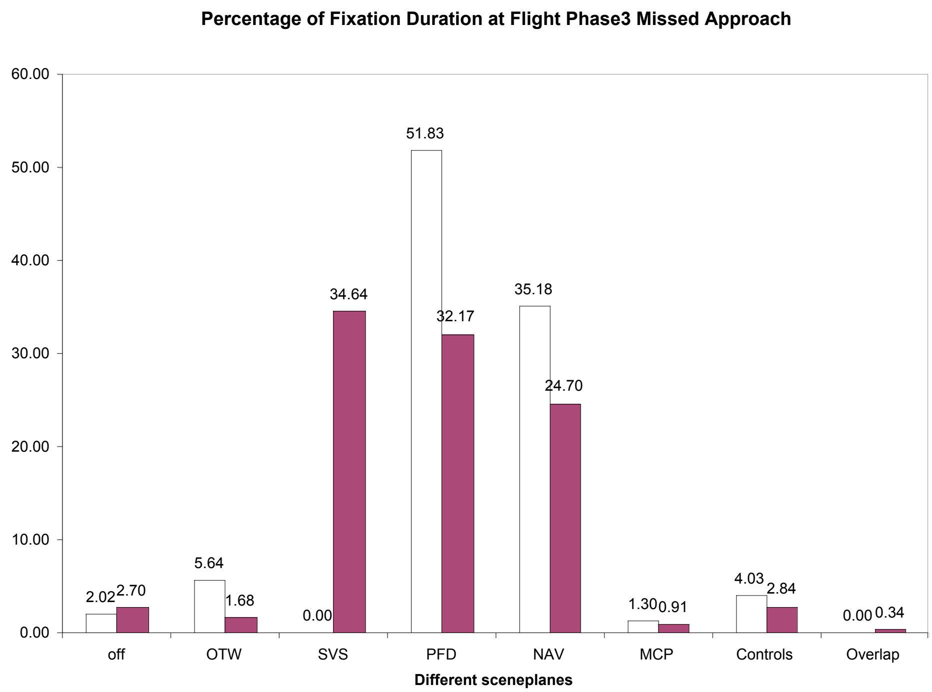


Figure 12. Percentage of fixation dwell time in baseline (light bars) and SVS (dark bars) conditions during flight phase 3 for the Missed Approach scenarios.

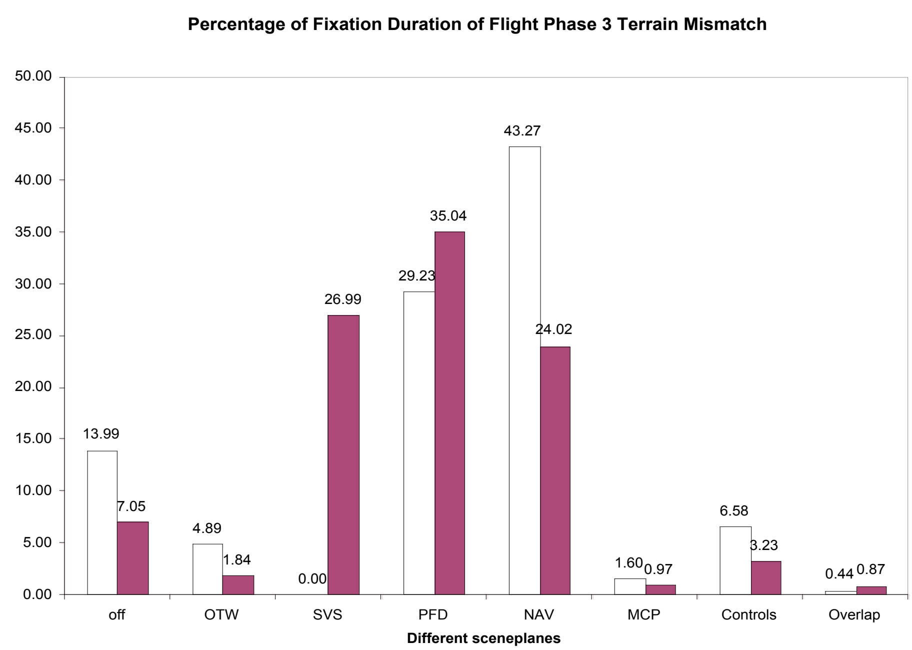


Figure 13. Percentage of fixation dwell time in baseline (light bars) and SVS (dark bars) conditions during flight phase 3 for the Terrain Mismatch scenarios.

Note how similar the OTW and SVS usage is for the three scenarios during phase 3. This is in stark contrast to phase 4, presented in Figures 14-16.

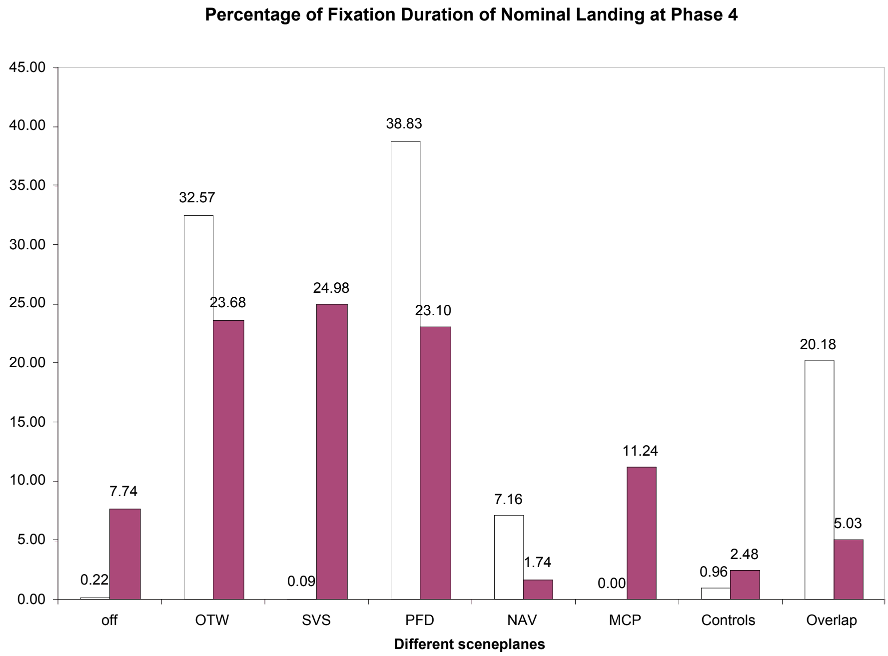


Figure 14. Percentage of fixation dwell time in baseline (light bars) and SVS (dark bars) conditions during flight phase 4 for the Nominal Approach scenarios.

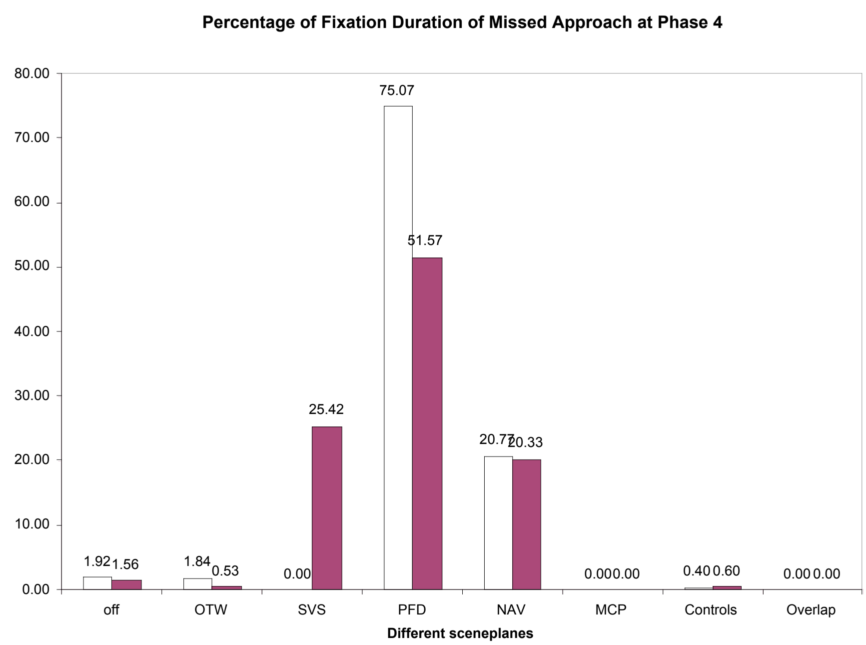


Figure 15. Percentage of fixation dwell time in baseline (light bars) and SVS (dark bars) conditions during flight phase 4 for the Missed Approach scenarios.

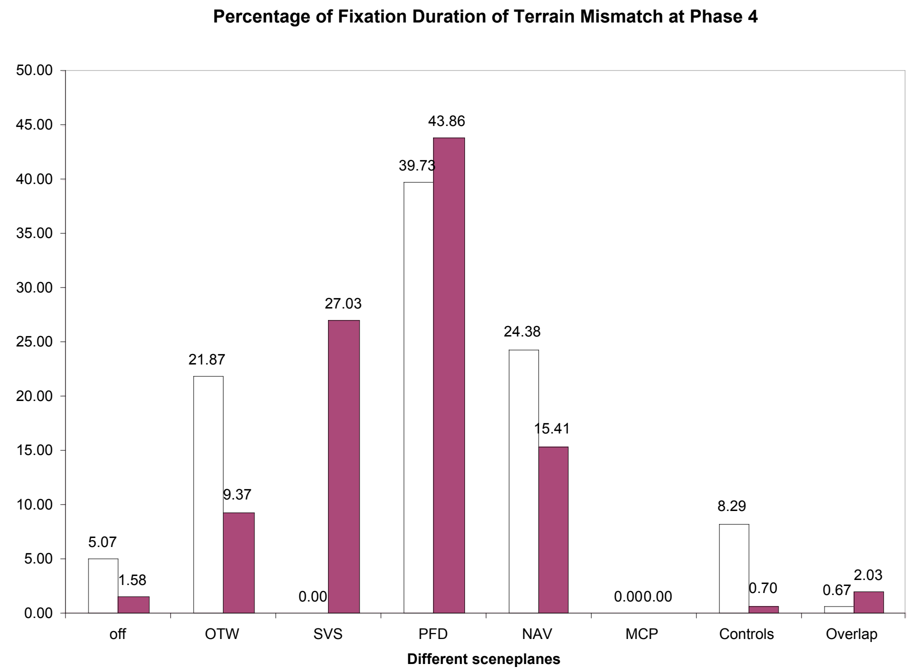


Figure 16. Percentage of fixation dwell time in baseline (light bars) and SVS (dark bars) conditions during flight phase 4 for the Terrain Mismatch scenarios.

The differences here are striking. In the nominal landing, pilots looked out the window quite a lot, even when the SVS was available. SVS use was certainly evident for the nominal approach in phase 4, but as seen in other graphs, this seems to come at least in part through less allocation of gaze to the PFD and possibly NAV. However, in the missed approach, OTW looking fell to virtually zero, both with and without SVS. Note that SVS use was roughly similar in nominal and missed approach, but in missed approach, this appears to have come almost entirely at the expense of gaze at the PFD.

For terrain mismatch conditions, the SVS use was again about the same (about a quarter of the total time), but this time it was *not* due to a reduction in looking at the PFD. Thus, use of the SVS is clearly conditioned on phase of flight and scenario. In addition, the SVS is evidently not simply a proxy for looking out the window. In particular, the SVS appears to serve many of the informational functions of the PFD, and perhaps also NAV, though to a lesser extent. This has other implications that will be discussed later.

The presence of the SVS may change not only how much the pilots look at various displays, but the strategies they use for acquiring information from their visual world. This should be reflected in changes to the order in which various pieces of information are acquired. Thus, we are in the process of expanding our data analysis to include sequential dependency information as well. That is, if fixation n is on the PFD, where is the most likely location for fixation $n+1$? Is this affected by the presence of the SVS, and if so, how? Since ACT-R produces a behavior stream, it should be possible to predict the various transition probabilities.

5. Description of Modeling Effort

A major focus of our modeling effort is to reproduce the trends observed on the basis of the detailed analyses presented in Section 4. We believe that to do so in a manner consistent with one of the major lessons learned from our Phase 1 research on taxi modeling, we must focus on the specification of cognition, the environment, and their interaction at a very fine grain of detail. We believe that this is especially crucial for modeling performance in a dynamic, interactive task. Highly detailed data and task analyses are critical, particularly for a cognitive architecture that operates at a very fine grain of temporal resolution, such as ACT-R. Thus, our focus has been on laying an appropriate foundation for the modeling effort. Our preference has been to eschew shortcuts for the promise of higher fidelity. While this has slowed certain aspects of our progress, we believe this will pay off later. Our approach has been to try to understand the major sources of both insight and constraint in generating our models. In addition to the detailed eye-movement analyses presented above, we identified three additional sources of information:

5.1 Task Analysis

Besides data analysis, our first order of business was to try to understand the task at a detailed level. This a challenge for this task because there is little overt action taken by the pilots; it appears to be primarily a supervisory control task, at least until the pilot takes manual control. However, the task is more complex than just that. To understand it, we have relied on three primary sources of information: the task analysis information collected and supplied by NASA Ames (Keller, Leiden, & Small, 2003); other related work in the human factors of aviation; and conversations with our subject matter expert (SME). We have synthesized these into the ACT-R formalism. An example of some of the resulting control structure appears in Figure 17.

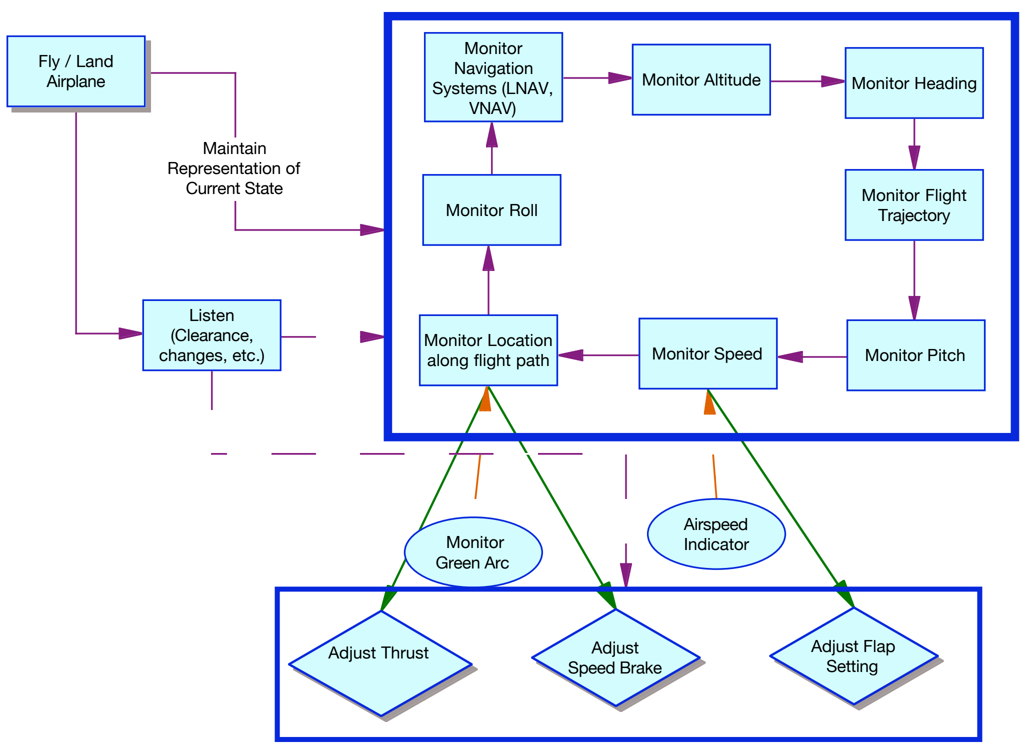


Figure 17. Flow of control resulting from task analyses.

The first insight from the knowledge engineering process is that the bulk of the task, particularly for the first two phases of flight, is primarily a monitoring task in which the pilot is engaged in maintaining his or her representation of the state of the aircraft. Additionally, we learned that pilots very actively check for a number of events and conditions which do not occur in the scenarios, such as late changes of wind direction that might lead to wind shear. Thus, for a lot of the time during the experimental trials, there is the appearance of little workload while in fact the pilots do still have a lot to do. This is fairly realistic for most landings which are, in fact, routine. However, pilots do have to monitor for non-routine conditions. In order to simulate the true workload accurately, we have included checks for many of these things in the model even though they do not occur in the scenarios.

5.2 ACT-R

The ACT-R architecture provides a great deal of constraint as well. Working within the parameters of the architecture sets certain boundaries and delimits scope. In particular, it means that we are modeling the task at a highly detailed level of analysis. ACT-R provides end-to-end modeling of the human operator side of the human-in-the-loop, from basic visual and auditory attentional operators to complex cognition and back down to basic motor movements. This impacts the strategies that are even possible and the way in which knowledge about dynamic state has to be updated to be maintained. A thorough review of ACT-R is far beyond the scope of this presentation. However, it should be noted that we are now using the most recent version of ACT-R, termed ACT-R 5.0 (see Anderson, Bothell, Byrne, & Lebiere, 2002 for a detailed description). ACT-R 5.0 incorporates the perceptual-motor extensions found in ACT-R/PM and provides for even more aggressive parallel execution of cognitive, perceptual, and motor operations than did the ACT-R/PM version of the system used for the taxi model work.

5.3 Extant Accounts of Relevant Phenomena

Because the eye-movement data are the primary focus of the modeling effort, we have examined other data and models in the “allocation of attention” domain in the human factors literature (e.g., Senders, 1964; Wickens, 2002). These are high-level (relative to ACT-R) accounts of how operators choose which objects to visually sample and at what frequency. The basic findings are that the rate at which particular displays are sampled depends jointly on the task importance of the displayed information as well as the rate of change of the information. As one might expect, more important information is sampled more often, and more dynamic information is sampled more often. We believe that these accounts provide a useful high-level starting point; we hope to provide the explanation for how these high-level phenomena emerge from a combination of task and environmental constraints and relatively low-level cognitive-perceptual capabilities. In other words, we recognize and have learned from theories that predict the percent of fixation time on particular information sources through mathematical modeling. Our goal, in contrast, is to create a process model from which such higher-level descriptions emerge as a function of the lower-level mechanisms in the model-environment system.

6. Focus and Intent of Modeling Effort

We had three major foci in the present effort:

6.1 A Dynamic, Closed-Loop Approach

One of the things which distinguishes an analysis at the level of a cognitive architecture such as ACT-R is that it is possible to “close the loop” of the human-machine system. That is, both the human and the evaluated system are modeled dynamically and in detail, and the two sub-models are coupled, yielding a model of the complete dynamic system. Work on the taxiing model revealed that fidelity of the machine/environment model was critical in understanding the performance of the human model; in particular, many of the "higher-level" decisions ultimately depended on "low-level" properties of the human-environment system. For example, the decision of "which strategy should I use to choose which direction to go?" often depended on things like the distance between the sign and the intersection as well as when the cognitive system was free to sample that part of the visual environment. Because ACT-R is fundamentally a non-linear system, small perturbations in the dynamic state of the human-environment system at one time can often lead to large differences in state or behavior further down the road.

Thus, we feel it is critical to continue with this rather complete, closed-loop approach. As previously mentioned, this means we have to contend with a great deal of detail in modeling the pilots' behavior, but ultimately we believe that path will lead to the best model.

6.2 An Adapted Pilot

Present efforts are based on modeling a pilot who is both knowledgeable about the task and well-adapted to it. We are neither modeling novice pilots nor the acquisition/development of piloting expertise. This limits the scope of the model but has other implications as well.

In particular, this means the task analysis information is, in some sense, "contaminated" by the fact that the pilots come into the task with a pre-existing strategy for how to sample the relevant displays. Because they know which information is most important and have a clear model of which information will be most dynamic, their strategies reflect this knowledge. That is, the relevance and rate of change for properties like altitude are known in advance by the pilots, so the pilot does not have to figure out how often to sample that information, he or she already knows how often it needs to be sampled. However, we believe that this has certain implications which we may want to relax later, see the section on later efforts for more details.

6.3 An Attention Allocation Focus

As mentioned previously, we believe the primary phenomenon to be explained here is how the pilots deploy their visual attention across the visual array and how this is (or is not) affected by the SVS. While this appears straightforward, there are some subtle issues here which we are exploring. For example, the ACT-R model produces time stamped individual shifts of visual attention (saccades) to small targets; we believe it is a mistake to attempt to map these directly to the individual saccades made by the pilots. Rather, such data can be analyzed at different levels of abstraction. For example, one could reasonably be interested only in more gross performance measures, such as the proportion of fixations on each scene plane, for which we have human data. We can run the model, which produces data at a much finer level of detail, but then extract these higher-level measures from the model run. In fact, this extraction can be performed with more or less the same set of analysis tools that were used on the human data.

An important research question is: What level of analysis is appropriate to guide design decisions? Did we want only the more gross measures such as proportion of fixations on each scene plane, or was it worthwhile to attempt to match the exact sequence of fixations generated by a model run with the exact sequence generated by one human trial? While the answer is somewhere in between, this is still an empirical question. Because ACT-R produces behavior at a fine grain size, we had the option of potentially examining behavior at multiple levels.

6.4 Implementation Approach

Many of the details of the implementation have already been discussed. The primary inputs to the cognitive model come from the task analysis; this is the source of the procedural knowledge and the bulk of the initial declarative knowledge given to ACT-R. The output of the model is a time stamped series of behaviors including individual attention shifts, speech output, button presses, and the like. The primary point of comparison for the model output is the human eye-tracking data, which can be examined at various levels of abstraction. One piece that has not been described in detail thus far is the other half of the simulation: the simulation of the aircraft. We have mocked up the primary displays (NAV, PFD, MCP, etc.) in the language of ACT-R so that it can directly "view" those pieces of the display. However, this is not enough; ACT-R requires a dynamic environment with which to interact. For instance, if the flap setting is changed by the model, there are certain expectations about downstream effects on flight performance. To make those happen properly, a simulation of the airplane is required. We have purchased the commercial software package X-Plane for this purpose and are in the process of linking X-Plane to ACT-R (note that X-Plane has been certified by the FAA for training pilots, see http://www.x-plane.com/FTD.html). Figure 18 presents a picture of X-Plane in action.

Figure 18. The X-Plane flight simulation package.

This linkage process is not trivial; we are writing a network interface (based on the UDP protocol) between the two programs from the ground up. X-Plane natively supports sending certain kinds of information such as altitude and heading via the network interface, but other things cannot be sent, including the view out the window. This represents something of a problem since the ACT-R model needs something to "see" out the window (and on the SVS). However, we believe this problem can be solved relatively straightforwardly by abstracting out only what the model would need to look for when it looks. For example, because we know the plane's absolute position and orientation with respect to the airport, we can determine whether whatever piece of information the model was seeking would be available. This task-oriented solution may have uses in other domains as well.

In addition, we have to supply X-Plane with the aircraft specifications (a 757) and the appropriate approach/navigation and FMC programming (e.g., fix points) for Santa Barbara. Fortunately, the 757 specifications and the airport and geography for Santa Barbara were freely available and could simply be plugged in. Figure 19 presents a diagram describing the system. System runs will involve initializing both ACT-R and X-Plane appropriately, running them, and collecting a trace of the output. X-Plane is designed to run in real time, so generating multiple simulation runs will be time-consuming. (However, there may be some workarounds for this and we are hoping to get X-Plane to run 2x or 4x real time.)

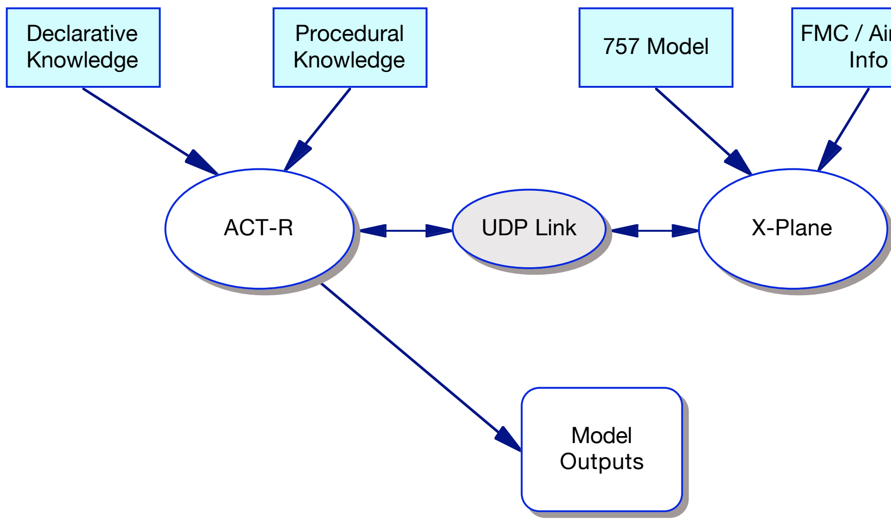


Figure 19. System overview.

7. Findings

7.1 Preliminary Results

Because the fully-coupled simulation is not yet completely operational, our findings are currently somewhat preliminary. However, we believe that we have still made substantial progress and, more importantly, gained significant insight. First, our initial data analysis shows that the SVS does indeed affect attention allocation, and that this is conditioned on phase of flight. Consider Figure 18, which shows the percentage of the dwell time by region of interest (ROI) for flight phase 1 (start to initial fix). Note the similarity between the non-SVS and SVS conditions. Contrast this with Figure 20, which presents the same data for phase 3 (final fix to decision altitude). Note how the pilots make little use of the SVS in phase 1, but in phase 3 their eyes are aimed at the SVS nearly a third of the time. As mentioned earlier, that the SVS is not simply a proxy for looking out the window in phase 3; pilots rarely look out the window at this phase. Instead, pilots look at the SVS and look less at the PFD and NAV displays.

At a high level, the model has a clear story for these data. The model predictions are based on the number of times a piece of information must be found and where the model will look for that piece of information. The model proportion presented here is simply the number of times attention will be directed to any particular display divided by the number of times attention will be directed to all relevant displays. When a piece of information could be found on the SVS as well as somewhere else (the PFD or OTW), the weak assumption was made that the model would look for that information from the SVS half the time and from the other source (PFD, OTW) the other half of the time.

Table 2 presents the overall (that is, not conditioned by phase of flight) data for both the human subjects and the model. This is essentially a static approximation of the dynamic ACT-R system, which may vary from this somewhat in final form. However, the initial analysis is encouraging; the fully-dynamic model should certainly be able to capture the patterns found in the data.

Table 2. Human data and analytic model predictions for fixation frequency

Region of Interest	*Data, no SVS*	*Model, no SVS*	*Data, with SVS*	*Model, with SVS*
NAV	0.39	0.30	0.28	0.27
PFD	0.38	0.44	0.33	0.30
MCP	0.07	0.19	0.03	0.20
OTW	0.03	0.07	0.03	0.03
SVS	-	-	0.21	0.20

What is important to note here is that the predictions for the SVS condition, in particular, are sensitive to local properties of the display. We currently use a rough estimate that if an item is available on the SVS then the model will look at the SVS half the time; this is currently a baseline assumption. In fact, if the model needs to look for a particular piece of information that is available in multiple locations (e.g., altitude, which is on both PFD and SVS), where it will look will be conditioned on where it is currently looking. ACT-R models are sensitive to local costs, and looking further away takes longer, so the model will prefer to look for the altitude on the SVS if it is already looking at the SVS.

Essentially, we see high-level properties of the model, such as its overall attention allocation behavior, as emergent from the combination of lower-level mechanisms and the structure of the task and environment. This should allow us to make predictions about even very small changes of the display; for instance, the model predicts that the overlay of airspeed and altitude on the SVS is a major factor in determining the degree to which the pilots will look at the SVS, as detailed below.

7.2 Implications

Given the previous section, some of the predicted implications for SVS design are fairly straightforward. For example, at the HFES conference this past October, a field test of an SVS system was described (Prinzel, et al., 2002). This SVS, however, was different from the SVS used in the Ames study for which we have data. Specifically, the SVS which was field tested had altitude and airspeed displayed in moving bars (the way they are displayed on the PFD) overlaid, which the modeled SVS does not. Our model would suggest that this is will lead to increased SVS usage because it makes rate of change of altitude and airspeed easier to obtain from the SVS. Our task analysis shows that the rate of change of altitude and airspeed are quite critical quantities at many points in the task, so making them available on the SVS should definitely lead to increased SVS usage.

However, this is not true for all the overlaid symbology on the SVS. For example, while it is true that pilots do need heading information, this is not as critical to place on the SVS because heading information generally does not change as fast, or is needed as often, as rate of change of altitude. On the flip side, there is other symbology that could be added that would likely lead to increased SVS use. The flight path predictor (FPP) is, we believe, a good bit of symbology, but we believe, based on the frequency with which the model needs to look at the NAV display, that this would be improved if the SVS also rendered the way points. Thus, flying to a way point would, (only in a visual sense), "simply" means keeping the FPP lined up with the way point indicator.

In general, the model predicts that the symbology overlaid on the SVS is a key component in determining how the SVS is integrated into the pilots' attention allocation strategies. This is almost certainly conditioned on phase of flight as well, suggesting perhaps that the symbology be configurable depending on phase of flight. Furthermore, we expect that the ACT-R model should be able to make predictions about the effects of adding or removing specific pieces of information.

We have gained other insights as well. First, even from the three subjects for whom data was provided by Ames, there were substantial individual differences, particularly at the more local levels. Because of this, and because such differences are likely to exist in the wider population of potential SVS users, we believe it would be a huge mistake to try to fit every aspect of these individuals' behaviors. Attempting to fit the complete scan path for any one subject would not only be laborious, it would almost certainly be an instance of fitting a great deal of noise. Just because the model is capable of generating fine-grained behavior does not mean that should be the basis of evaluation; rather, we believe more abstracted measures will do a better job of "smoothing out" individual difference noise and thus should constitute the model's criteria. We are not yet certain exactly what the best measures should be, but we believe this is an important question that we likely would not have considered without the combination of our model and the data we have in hand.

8. Progress and Lesson Learned

8.1 Progress and Advances

While there is still much more work to be done and many things to learn, we believe we have generated several advances. First, the model is not tied to any of the specifics of the scenario. If the FMC is pre-programmed correctly and the model is given a relatively modest amount of knowledge about the airport, the model should be able to run through the approach fixes for any approach and landing scenario, as long as no serious maneuvers are required. This could potentially be a win for future aviation safety research. Second, the network interface we are developing could have wide applicability, as many simulation environments (e.g. video games) use similar communication protocols; this may make it possible to connect ACT-R to a wider range of environments. This should be particularly powerful when combined with the task-oriented solution we have generated to the out-the-window vision problem.

In addition, we believe that we may have some leverage on some other high-level and abstract human factors constructs, such as "situation awareness." There is no box or section of the ACT-R architecture that one could point to as being situation awareness. Rather, we have observed that the model has to keep a number of pieces of information available at various times (some things, like altitude, all the time); the accessibility of the set of needed information about the aircraft's state might be termed the model's situation awareness, but it is not a unitary thing. It is both distributed, in that it lives in multiple declarative memory elements, and dynamic, in that different pieces are needed and "refreshed" by checking the environment at different rates. We hope ultimately that this work will lead toward more formal definitions, at least in an ACT-R context, of a number of terms from the human factors literature (e.g., situation awareness, workload) that are currently somewhat vague.

Furthermore, we are excited by the idea that previous results in the attention allocation area might be explained by lower-level mechanisms in our ACT-R framework. For instance, consider the effects of rate of change of a display item on the human sampling rate. We believe this effect falls very naturally out of ACT-R's memory system. When the model looks at the display for a particular piece of information, a representation of that information is created in declarative memory. However, ACT-R's memory decays over time, which creates a need to re-sample the environment. If the environment is re-sampled and yields the same result, rather than creating a new representation of say, airspeed, the activation of the extant representation will be incremented. This means it will take longer for that piece of information to decay, which means it will not have to be sampled as often. Thus, information that changes slowly will more often yield the same value when sampled, and thus will decay more slowly, requiring less frequent sampling.

8.2 Challenges Remaining

Doing a detailed simulation of human-in-the-loop performance in a domain this complex is fraught with challenges; many of them have already been described. Probably the biggest thing that could have gone more smoothly and should be considered for future efforts is to give the modeling teams direct access to the simulator code; the X-Plane solution we believe will ultimately work, but it has been slow going. On the positive side, there should be a downstream payoff for future efforts to link ACT-R to other systems.

Nonetheless, we realize sometimes there are limitations of time and energy for what can be provided to modeling teams, and it is clear that the rich eye-tracking data is more important than providing such linkages.

8.3 Future Directions

Obviously, there is still a great deal of work to be done to completely "close the loop;" this is currently our top priority. Once that is done, we hope to explore the design space for the SVS a little, and will try variants of the current SVS symbology to assess their impact on the model's performance. We are hoping this will lead to greater insight into the evaluation of SVS technology.

In addition, we would like to explore "de-adapting" the task analysis. One of the issues with many task analyses as they currently stand is that they include the operator's attunement to the constraints of the environment and may not be terribly useful at predicting how performance would be if the environment were different. We hope to produce a more abstract model, possibly more complex than is needed to mimic, from an input-output perspective, the over-learned routines that underlie skilled adaptation to an existing cockpit design. To us at least, a more abstract model seems to be necessary to allow us to predict the cognitive implications of novel changes to the cockpit design from a first-principles perspective.

9. References

Anderson, J. R., Bothell, D., Byrne, M. D., & Lebiere, C. (2002). An integrated theory of the mind. Manuscript sumitted to *Psychological Reivew* and available at http://actr.psy.cmu.edu/papers/403/IntegratedTheory.pdf.

Bowers, C. A. and Jentsch, F. (2001). Use of commercial, off-the-shelf, simulations for team research. In E. Salas (Ed.), *Advances in Human Performance and Cognitive Engineering Research, Vol. 1*. Mahwah, NJ: Erlbuam. (pp. 293-318).

Goodman, A., Hooey, B.L. Foyle, D.C., & Wilson, J.R. (2003). Characterizing visual performance during approach and landing with and without a synthetic vision display: A part task study. *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.

Foyle, D.C., Goodman, A., & Hooey, B.L. (2003). An overview of the NASA Aviation Safety Program (AvSP) System-Wide Accident Prevention (SWAP) Human Performance Modeling Element. *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., Small. R. (2003). Cognitive task analysis of commercial jet aircraft pilots during instrument approaches for baseline and synthetic vision displays. In, 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.

Prinzel, L. J., Kramer, L. J., Comstock, J. R., Bailey, R. E., Hughes, M. F., Parrish, R. V. (2002). NASA Synthetic Vision EGE Flight Test. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 135-139). Santa Monica, CA: Human Factors and Ergonomics Society.

Senders, J. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Transactions on Human Factors in Electronics, HFE-5*, 2-6.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science*. (pp 1-19).

Human Performance Modeling Predictions in Reduced Visibility Operation With and Without the Use of Synthetic Vision System Operations

Brian F. Gore, Savita Verma, Kevin M. Corker, Amit Jadhav, and Eromi Guneratne
San Jose State University
San Jose, California

Abstract
The San Jose State University human performance modeling team undertook this human performance modeling research effort to predict the performance of operators using the Synthetic Vision System (SVS), with support of the NASA Aviation Safety Program. Test scenarios were developed and procedures were established based on the NASA Ames part-task human-in-the-loop simulation for both the baseline (current technology) operations and the advanced SVS operations conditions. The aircraft performed approaches to landing at Santa Barbara Airport flying under Instrument Meteorological Conditions (IMC) with "current day" technologies or "future" cockpit configuration (SVS display). The Air MIDAS model was augmented to handle the flight procedures observed in the human-in-the-loop simulation. The standard Air MIDAS model of visual performance was augmented to include the affect of contrast legibility and visual search/reading time to account for performance using the synthetic vision system. After model development, a simulation test was run on approach under conditions of baseline, SVS, and SVS with sidestep maneuver required. High correlations were found between the modeled procedures and information-seeking behavior and that of the human operator's performance in simulation. The model's data were subjected to verification and validation analyses.

1. Introduction
NASA is developing a number of technologies designed to aid the flight crew in the safe operation of the aircraft under conditions that in the past have been shown to contribute to increased hazards in aviation operations. Those technologies have a common purpose in aiding the flight crew by providing information that has either been not available (e.g. improved traffic position information or rapid update of local meteorological conditions like turbulence) or has been obscured and degraded (e.g. visual acuity reduction in weather and at night). The advancements in computational techniques, sensor and communication technologies have resulted in an enviable design situation in which the amount and quality of information available is large and therefore must be carefully selected to avoid overwhelming the flight crew. Interesting issues of information selection, information integration requirements and display operation are open to investigation in the conceptual and early design stages of the systems development.

1.1 Synthetic Vision System
Recently, NASA has been developing augmentative technologies comprising a synthetic vision system (SVS) for commercial aviation as well as for business jets, and general aviation operations. The system is designed to generate a texture-mapped (or wire-frame) display of the terrain in proximity to the aircraft. Text and other symbology will be overlaid onto the terrain display to display, for instance, the aircraft itself, its velocity, a "follow-me" aircraft, a "tunnel-in-the-sky" indication of the route, and indications of other nearby aircraft. In addition, flight controls (air speed, attitude, pitch, etc.) will be overlaid on the display. A more complete review of the several designs under development for the support and provision of synthetic vision can be found in Corker & Guneratne (2002). In addition, the existing display elements of current aircraft will be maintained in an SVS-equipped aircraft. Providing both of these sources of information may be problematic. On one hand they support cross checking of flight deck systems, on the other hand two sources of information that are similar in source and content, but different in presentation mode may cause transformation workload for the pilot. When systems such as the one being proposed for the SVS are being designed, we suggest that, in early design phases, computational human performance models can be used to predict various performance effects of introducing such augmented technologies.

2. Human Performance Modeling

The use of the human performance modeling methodology has been suggested as an effective means to study concepts in complex systems or those designs that are very early in their design phases (National Academy Press, 1990). In the type of human performance modeling undertaken in this study, the parameters of human behavior embedded within the model framework are based on empirical research in both basic and applied human performance. The modeled operator is then set to interact with computer-generated representations of the operating environment over a series of repeated runs in much the same manner as testing human subjects over repeated experimental sessions. Elements of the human performance model (for example, performance time for a particular task) can be made a stochastic variable and their values can fluctuate across these multiple runs. The model of human performance enables predictions of behavior based on elementary perception, attention, working memory (WM), long-term memory (LTM) and decision-making models of human behaviors. This modeling approach focuses on micro models of human performance that feed-forward and feedback to other constituent models in the human system depending on the context and on mission needs and requirements.

Human performance models have produced validated predictions of human performance within complex operating environments ranging from highly advanced military systems (Atencio, 1998; 1994), nuclear power plan operations (Corker, 1994), and advanced concepts in aviation (Corker, Gore, Kennedy & Lane 2000). In this study, the human performance modeling software tool, Air Man-machine Integration Design and Analysis System (MIDAS), was used to generate predictions of human performance using the synthetic vision system (SVS).

2.1 Air MIDAS

The Air MIDAS software (a NASA Ames Research Center, San Jose State University development effort) is a performance prediction software tool that uses models of human performance within an integrated computational framework to generate workload, and activity timelines in response to operational environments (Gore, 2002).[1] The main components of the model exercised in this study were the simulated operator's world representation, and the symbolic operator model (SOM) representing perceptual and cognitive activities of an agent. In the SOM, the Updateable World Representation (UWR) contains information about the environment, crew-station, vehicle, physical constraints and the terrain. Updates of the states of these elements are provided through the perceptual and attention processes of the SOM. The world representation serves to trigger activities in the simulated operator to serve mission goals or respond to anomalies. The UWR also contains the WM of the simulated operator, the domain knowledge, and a goal-based procedural activity structure. Activities to be performed are managed through a queuing process and scheduled according to priority and resource availability. Four resource pools (Visual, Auditory, Cognitive, and Psychomotor) are checked for resource availability in response to the demands for those resources by the required tasks (McCracken & Aldrich, 1984). Figure 1 outlines the model organization and flow pattern associated with information entering into the modeled operator.

[1] For a complete review of the Air MIDAS model see Corker (2000).

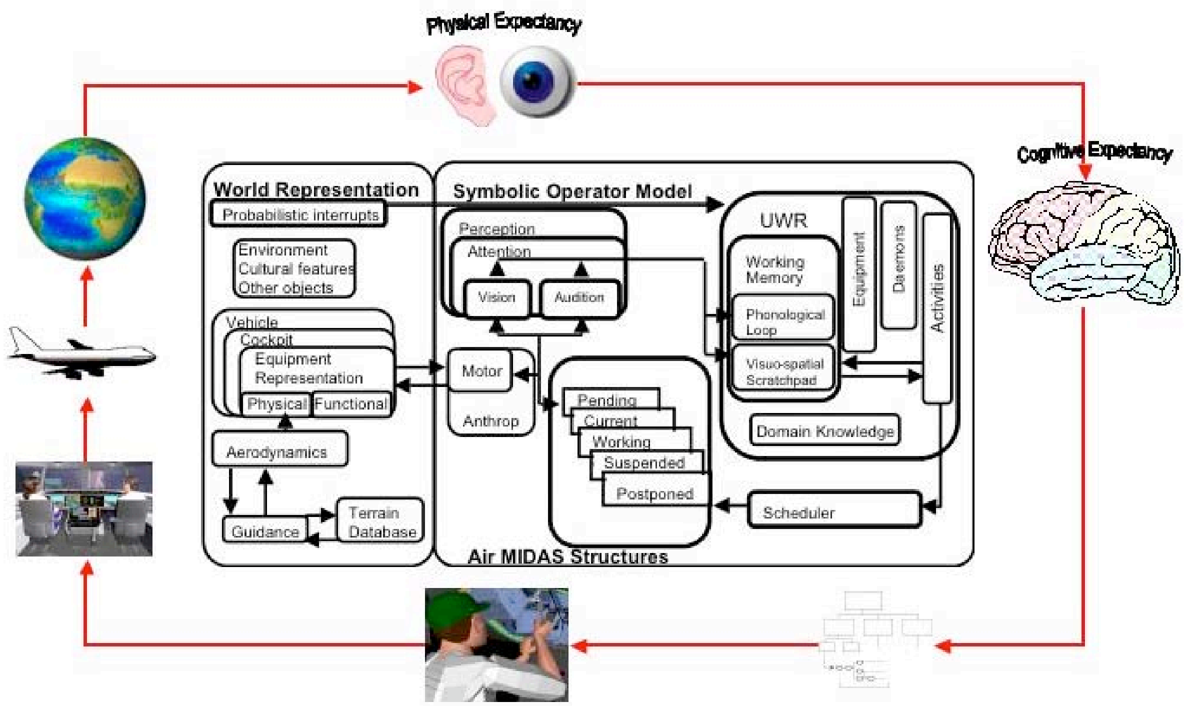


Figure 2. Air MIDAS Structure and Control Flow (from Gore, 2003).

Visual information such as that provided by the SVS or the out the window information is perceived and attended by the Air MIDAS operator. This external information is passed into the Symbolic Operator Model (SOM) through its attention and perception models. Once this information is perceived, it is passed into the Updateable World Representation (UWR) structure that contains the WM, phonological loop, a visuo-spatial scratchpad, rules for invoking and retaining memory information and the domain knowledge of the condition surrounding the operator.

In the SVS example, the operator perceives, for instance, descent-related information either from instruments or from the out-the-window view. These data trigger a series of rules to satisfy flight goals. In the case under study here, perceptual processes associated with the SVS system and/or the out-the-window observation are critically important. Their development is described below.

2.2 Air MIDAS Visual Perception Model

Perception in the Air MIDAS model proceeds by initiation of a visual activity (e.g. scan-instruments) that updates information in MIDAS every 'tick' (a 100 ms time increment). In support of the SVS experiment, the Air MIDAS perceptual process was enhanced.

In keeping with an approach to landing under both visual and instrument meteorological conditions, the Air MIDAS perceptual functions were developed to include in-cockpit scanning, both with and without the display augmentation of the SVS system. In addition, we included an out-the-window visual capture model for detecting features, aircraft stability, heading and position, associated with decisions on approach. The equipment representation and visual perception were refined to include the behavior of a human operator interacting with display technologies on which various electronic visual enhancements (e.g. runway center and sidelines) were contained.

2.3 Visual Perception Process

The human-in-the-loop eye fixation data for the baseline condition was compared with the model eye fixation for the baseline condition. The scan patterns of baseline condition were created using data adapted from Mumaw et al. (2000). The human-in-the-loop eye fixation data for the SVS condition was compared with the fixation data on the same scenario generated by the model. In generating the predictions of fixation pattern in the MIDAS software, it was expected that the SVS display would replace the out-the-world fixation times and that fixation on the SVS display would be followed by fixation on the PFD. To incorporate the SVS into the scan pattern of the flight crew, the OTW percentages were replaced by SVS scans. A description of each of the individual components as numbered in Figure 2 will be described below.

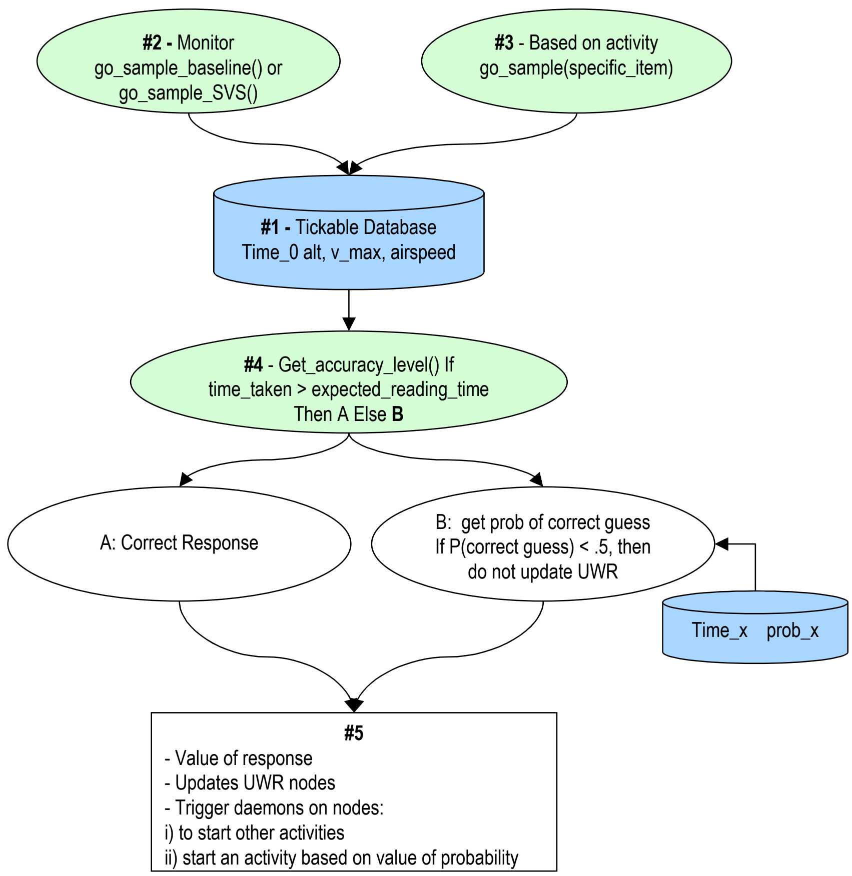


Figure 3. Flow of information for Vision Model.[2]

2.3.1 Component #1 – Flight Information Database

The database was designed to follow an incremental 100 ms update sequencing as visual perception activities are performed. Equipment components were incorporated in the simulation to provide the Air MIDAS operator with required flight-related information. These included the Primary Flight Display (PFD), the Mode Control Panel (MCP), the Navigation Display (ND), the out the window

[2] Components #1 through #5 are described below.

scene, the Synthetic Vision System (SVS), and other flight control information such as the flaps, the throttle and the speed brake controls.

A database was created to provide a shared database between the two Air MIDAS flight crew operators. The PFD provided the Air MIDAS operator with altitude, airspeed, attitude, heading, and the flight mode enunciator. The MCP provided the vertical speed, the MCP altitude and the MCP altitude dial. The ND information provided the aircraft true heading information. The OTW provided the simulated world landing information. The SVS provided the augmented display technology to enhance flight crew situation awareness on approach in instrument landing conditions. Some software daemons (responding to critical parameter values) were also incorporated based on the UWR nodes that got triggered.

2.3.2 Component #2 – Monitor – "Go_Sample" Baseline

As represented in Figure 2 through the monitoring node and Figure 3 through the detailed description of the go-sample structure, information from the database flows into the Air MIDAS operator through the augmented visual system. The baseline scan pattern was developed and used for the non-SVS scenario. Modifications were made to the scan pattern for the SVS scenario. Figure 3 represents the information flow.

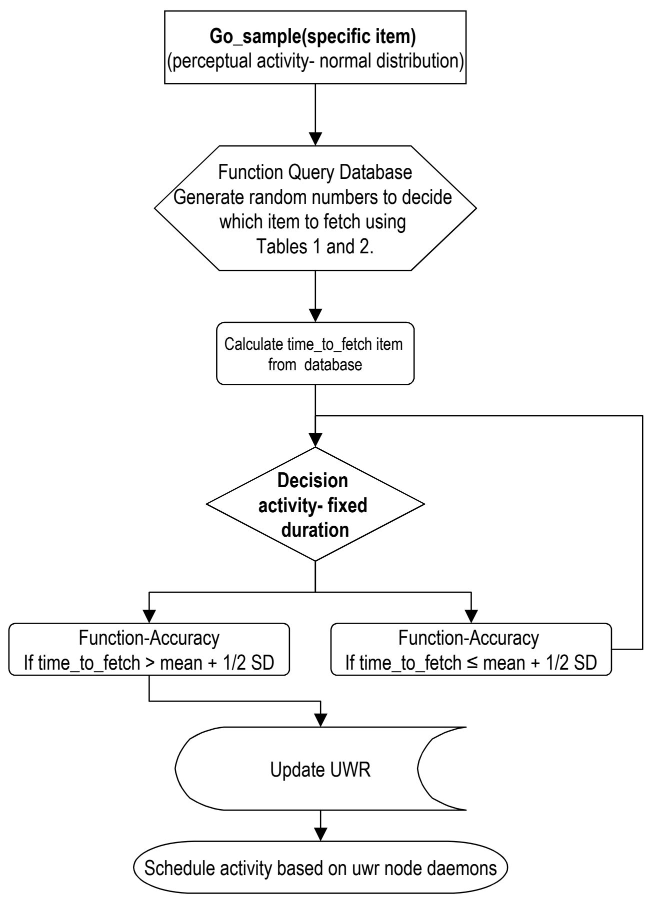


Figure 4. Flow of information when scan patterns are implemented.

The normal internal scan pattern and dwell time was based on NASA's report on the Analysis of Pilot's Monitoring and Performance on Highly Automated Flight Decks generated by Mumaw et al., (2000). These data can be found in Tables 1 and 2 below.

Table 2. Internal Dwell Time Percentages and Locations during VNAV Descent.

	Percentage Dwell	**Proportion**	Mean Dwell Duration (sec)
PFD	32% (32/82)	**0.39**	0.68
ND	33%	**0.40**	1.75
MCP	3%	**0.04**	0.72
Out of Window	1%	**0.01**	1.38
Other	13%	**0.16**	2.00
Total	82%		

Table 3. PFD Area of Interest (AOI) Percent Dwell Time for VNAV descent.

	Percentage Dwell	**Proportion**	Mean Dwell Duration
PFD airspeed	22%	**0.27**	0.68
PFD attitude	28%	**0.34**	0.54
PFD altitude	24%	**0.29**	0.59
PFD heading	3%	**0.04**	0.44
PFD FMAs	5%	**0.06**	0.41

The Air MIDAS model operates according to a sampling activity to obtain the information from the world. The gaze location is based on the proportions from Tables 1 and 2 above. The sampling activity for the SVS involved a scan of different equipment components. These included the third component called the "go-sample (specific target)" search pattern.

2.3.3 Component #3 - Details on go_sample (specific target)

Data from eye movement planning research suggests that humans perform a sample activity whenever an activity demands it. The "go_sample (specific target)" activity gets triggered whenever an activity requires current visually-provided information for its performance. Figure 4 below demonstrates the flow of information for the "go_sample" activity.

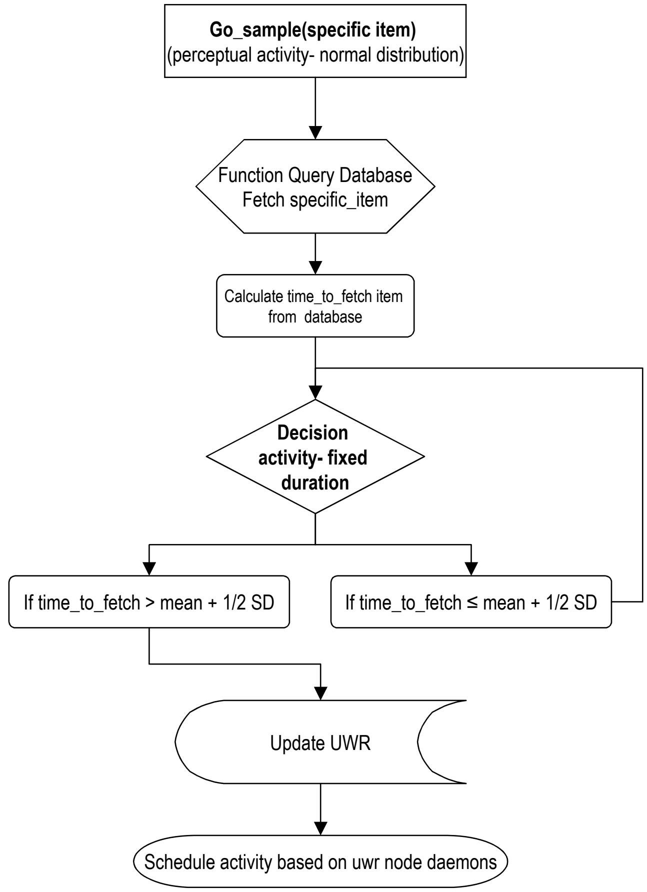


Figure 5. Flow of information when Go-sample (specific_item) is implemented.

2.3.4 Component #4 - Details on Accuracy Function

Information flow/sequence is only one part of the visual system that needs to be understood when explaining the process behind information uptake into the model representation. There are two other principal functions at work in information uptake, the expected reading rate and the accuracy functions.

2.3.4.1 Expected Reading Rate

The fixations associated with reading information from normal instruments should last a minimum of 200 ms (Landy, 2002). Visual fixations that move from an inside fixation to an outside fixation were increased by 500 ms to adjust for accommodation in the expected reading time. These considerations provide a reading time per character of 244 ms.

2.3.4.2 Accuracy Function

Table 3 (below) provides the mean, SD, min and max values of dwell duration for the Vertical Navigation portion of the descent phase of flight. This phase of flight possesses certain characteristic and required scan patterns and information searching behaviors of the flight crew's visual system. The data outlined below were used for building the accuracy function within the model's visual system.

Table 4. VNAV Descent Phase Data.

	MEAN	SD	MAX	MIN
Off AOIs	2.00	0.89	3.61	0.72
Out Win.	1.39	0.70	3.38	0.62
MCP	0.72	0.20	1.09	0.35
ND	1.75	0.47	3.08	1.16
CDU	1.63	0.60	2.67	0.80
PFD	0.68	0.11	0.90	0.50
PFD-ATT	0.40	0.26	0.82	0.00
PFD-Roll	0.44	0.33	1.34	0.00
PFD-Pitch	0.37	0.28	1.15	0.00
PFD-AS	0.68	0.40	1.41	0.00
PFD-ADI	0.54	0.29	1.04	0.00
PFD-ALT	0.59	0.31	0.90	0.00
PFD-HDG	0.44	0.27	1.08	0.00

Table 3 outlines the time associated with the situations when the visual scan is: (i) not fixated on anything in the cockpit or outside of the cockpit (Off AOI), (ii) when the fixation is out the window (Out Win), (iii) when the fixation on the MCP, (iv) when the fixation on the ND, (v) when the fixation on the Control Display Unit, and (vi) when the fixation on the various PFD readings. The decision to update the world representation of the Air MIDAS operator ("update_uwr") is based on the logic that states "if time_to_fetch item > Mean + 1/2 SD then update uwr, else do not update uwr". The assumptions associated with this exponential function can be found in Figure 5.

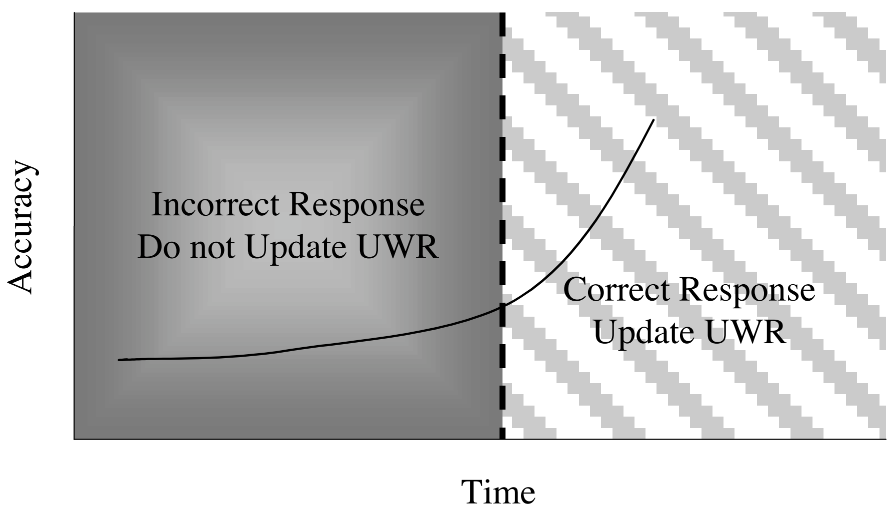


Figure 6. Accuracy function assumed as an increasing exponential function.

3. Simulation Experiment

An experiment was conducted to evaluate the impact of the synthetic vision system on information seeking and on flight procedures in an approach to landing both with and without the SVS system.

3.1 Participants

No human subjects were used in the current Human Performance Modeling simulation project. Human performance data came from the prior part-task simulation of the NASA HPM Organizing Team (Goodman, Hooey, Foyle, & Wilson, 2003). All perceptual model data came from either existing micro models within Air MIDAS (visual perception model - Remington, Johnston & Yantis, 1992; visual processing and field of view information – Arditi and Azueta (1992); Lubin and Bergen (1992) or from research conducted by Landy (2002). All procedural timing data came from tables of human performance load values based on the McCracken and Aldrich scales of procedural performance loads (McCracken & Aldrich, 1984), procedural specifications came from discussions with Subject Matter Experts (SMEs) provided by the HPM Organizing Team. Two flight crewmembers were modeled in this effort, the captain and the first officer and 5 runs were completed for each of the blocks as per the research design denoted in Table 4 below.

3.2 Apparatus

This computational human performance model, Air MIDAS, generated predictions of the operator's performance with the SVS technologies being introduced into the cockpit. Air MIDAS operates on a SGI IRIX 6.2 platform on a SGI Indy (R5000) workstation containing 96 Megabytes of Random Access Memory (RAM). Air MIDAS also operates on a Windows NT platform with minimum 96 MB of RAM.

3.3 Procedure

The experiment was conducted using the design illustrated in table 4. Based on the NASA HPM Organizing Team's prior simulation runs, we selected three of the scenarios (denoted in **bold** in Table 4 below) to exercise our model.

Table 5. NASA SVS simulation variables: Bold denotes Team HAIL scenarios modeled.

		Current Day	**Current Day Display**	**Future SVS Display**
		Good (VMC)	**Low Visibility (IMC)**	**Low Visibility (IMC)**
Approach Event	**Nominal Approach (nominal landing)**	*Scenario #1*	**Scenario #4**	**Scenario #7**
	Late Reassignment (side-step & land)	*Scenario #2*		**Scenario #8**
	Missed Approach (go-around)	*Scenario #3*	*Scenario #5*	*Scenario #9*
	Terrain Mismatch (go-around)		*Scenario #6*	*Scenario #10*

The nominal approach refers to the normal aircraft descent approach pattern with no deviation in flight plan to the runway surface. The late re-assignment approach refers to a request by Air Traffic Control for the aircraft to modify the approach plan and land the aircraft on a parallel runway. The current-day display refers to the current cockpit configuration when the aircraft is on the approach and landing phase of flight. The future SVS display refers to the display augmentations and resultant procedural changes associated with SVS operation. The visibility classification of either "good" or "low" refers to the degree to which the flight deck could see the external environment and the runway. Good visibility meant no visibility limitations while the low visibility meant no visibility until the aircraft broke through the cloud cover at 800 feet, very close to the existing minimum altitude decision height required for aircraft landing. The scenario numbers are the scenario numbers that were utilized by the NASA Organizing Team as a means of identifying the scenario for appropriate data collection.[3]

Figure 6 below outlines the RNAV (GPS) aircraft approach path, the altitude relative to the runway, the environment, and the decisions and responses necessary for safely landing a Boeing 757 aircraft during nominal operations. The RNAV procedures that were generated for the SJSU HPM simulation used information provided by the NASA Organizing Team as well as information from Boeing Subject Matter Experts (SMEs). These baseline procedures were then modified to include the SVS within the internal scan process of the flight crew. The performance of the modeled flight crew was measured in terms of event sequences, fixation patterns and workload estimates. The performance of most interest here is the performance of the model compared the performance of the human-in-the-loop simulation tests.

Rules to guide model behavior were developed based on the procedures required for approach and landing. These rules/priorities were:

1. Altitude information update is a priority in information seeking,
2. 500 foot altitude is always signaled and all scans below 500 feet are always an out the window scans,
3. Crosscheck between the crew-members always occurs.

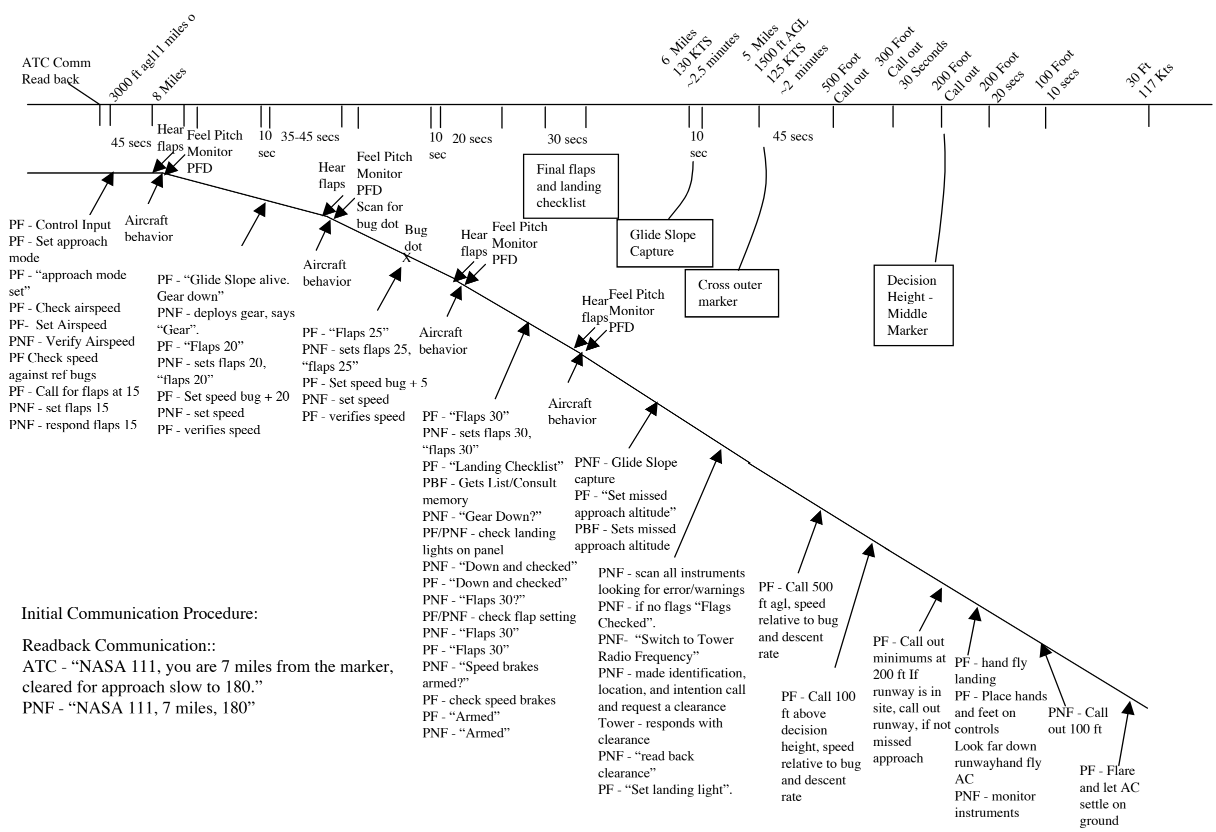


Figure 7. Procedural Sequence as Aircraft Approaches Santa Barbara Airport.

[3] For a more detailed description of the scenarios completed in the NASA SVS Simulation please consult Goodman, Hooey, Foyle, & Wilson (2003).

3.4 Data Collection

Data were collected at 100 Hz and post-processed by mapping to the event sequences in the simulation. The data were collected from the Final Approach Fix (FAF) to just before aircraft touch down. The data of interest were those associated with the point of aircraft break out, crew response time to the information in the simulated environment, and the procedural sequences associated with descent. The model was run under normal and low visibility conditions, both with and without the SVS, and either requiring or not requiring a sidestep maneuver. Five simulation runs were completed for each of the scenarios.

3.5 Results

The baseline runs served two purposes. First, we were interested in assuring that the model's operation produced data consistent with human performance in a baseline model, verification; and, second, that the verified baseline produced data that was predictive of human performance under new operational conditions, validation (Law & Kelton, 2000; Balci, 1998).

3.5.1 Fixation Frequency Analysis

Verification simulation trials on approach under (i) baseline without SVS, (ii) baseline with SVS and (iii) sidestep with SVS were conducted. As the data against which the verification was conducted did not reflect SVS use, assumptions as to the informational equivalence of out-the-window information seeking and SVS use were made. A strong correlation was found between the Mumaw et al.'s (2000) percent of fixations data and the Air MIDAS data across all scenarios. The correlation coefficients are as follows: (i) baseline (without SVS and with direct comparison to the source data) r^2 = 0.9936l; (ii) operation with SVS r^2 = 0.9955; (iii) SVS with sidestep r^2 = 0.9948. Figures 7 and 8 demonstrate the respective elements within the flight crew agent's scan pattern of the crew station and external environment. These data indicate that the procedural and visual sampling behavior largely replicate the source data of human performance. This is verification that the model behaves as designed and doesn't corrupt the seed human performance data.

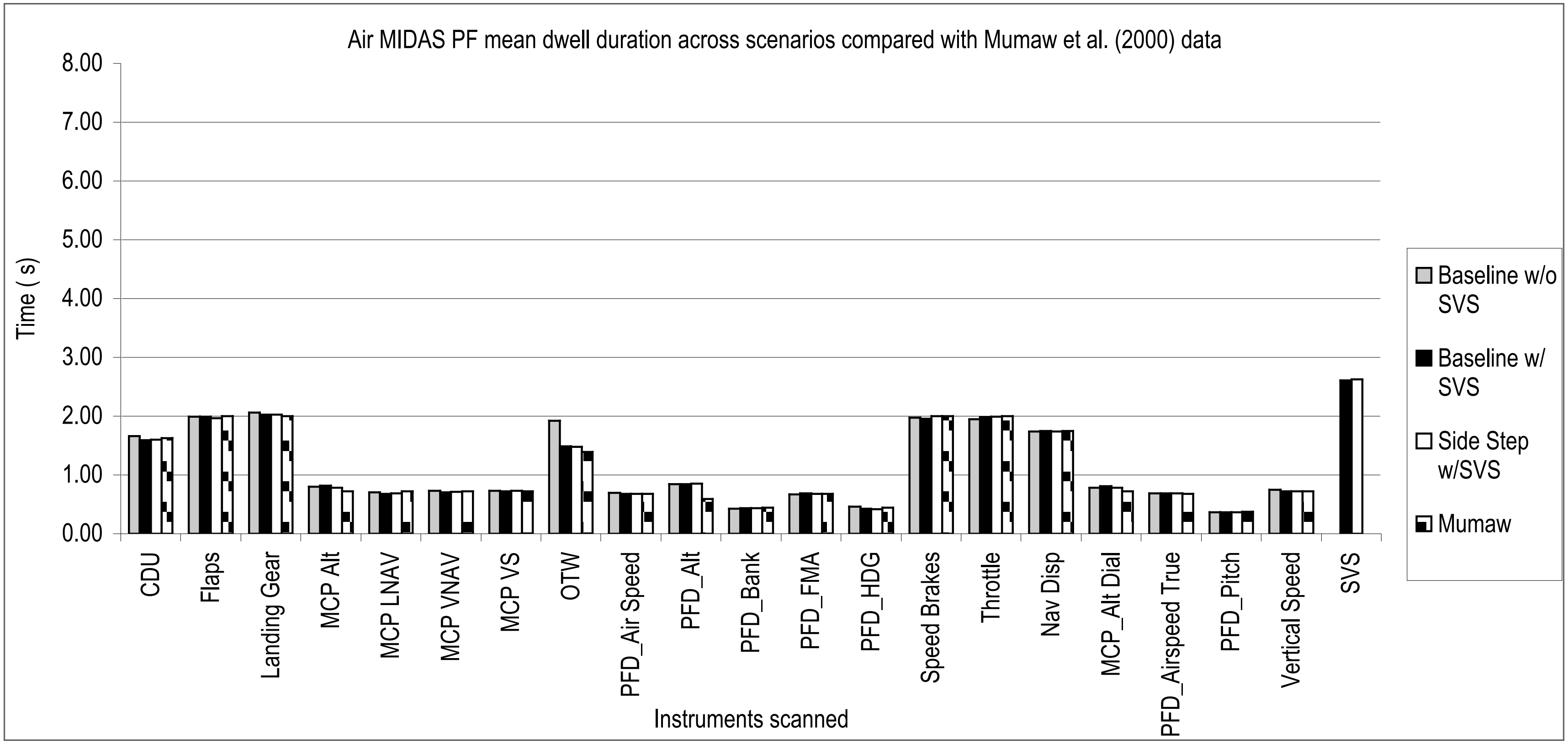


Figure 8. Air MIDAS mean Pilot Flying (PF) dwell duration compared with Mumaw et al. (2000) HITL data across scenarios.

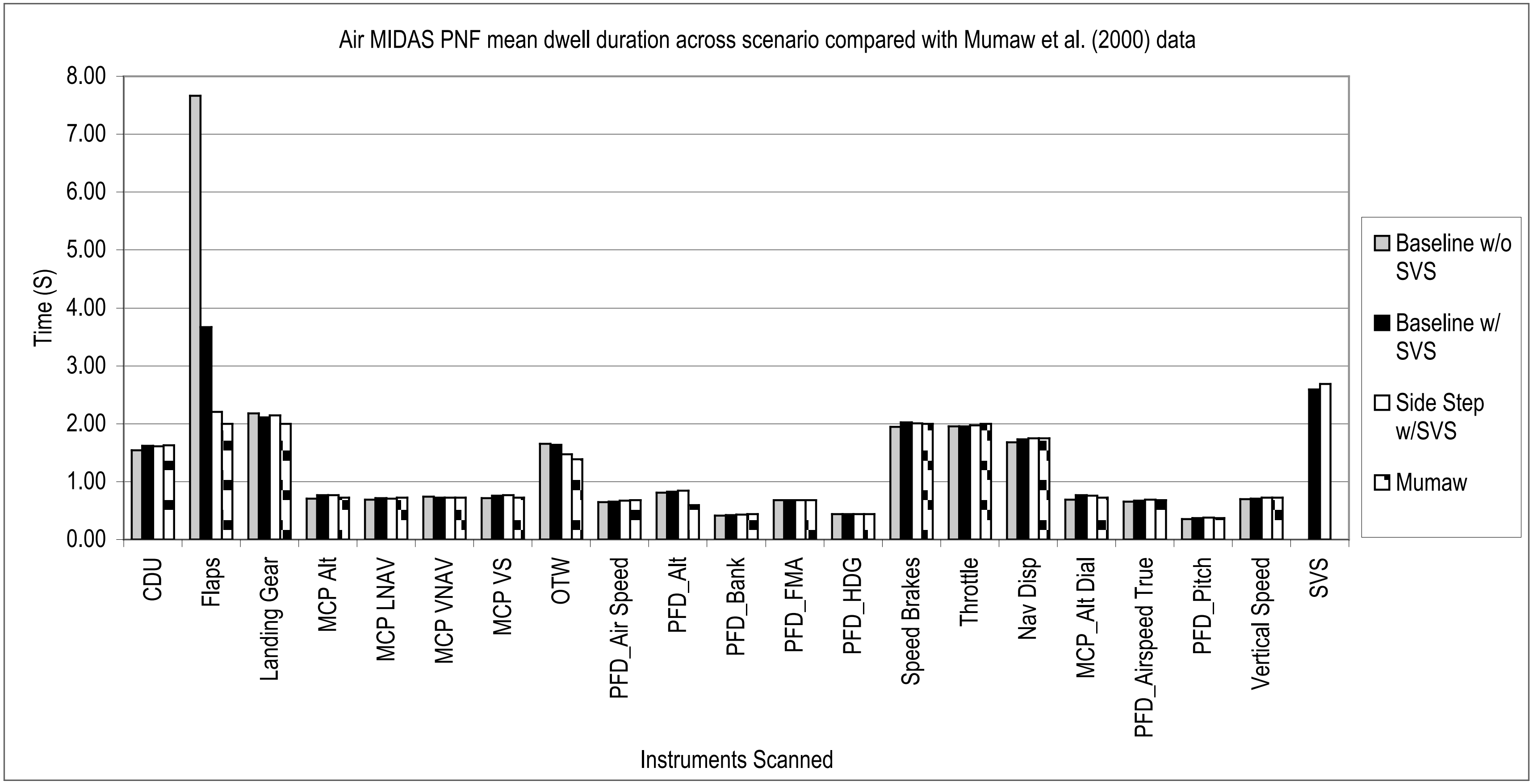


Figure 9. Air MIDAS mean Pilot Not Flying (PNF) dwell duration compared with Mumaw et al. (2000) HITL data across scenarios.

The predictive validity of the Air MIDAS model was also tested by running the model through three simulation conditions based on those undertaken in the NASA part-task experiment. Validation of the model's visual scanning behavior was examined by comparing model-generated dwell frequency to the human flight crew dwell frequency. The correlations between the NASA part-task simulation and the Air MIDAS data are as follows: (i) baseline $r^2 = 0.7608$; (ii) with SVS operation $r^2 = 0.\ 8782$; and (iii) SVS with sidestep $r^2 = 0.\ 5538$. An examination of each of the respective model-human dwell percentage locations comparisons by scenario can be found in the following three figures.

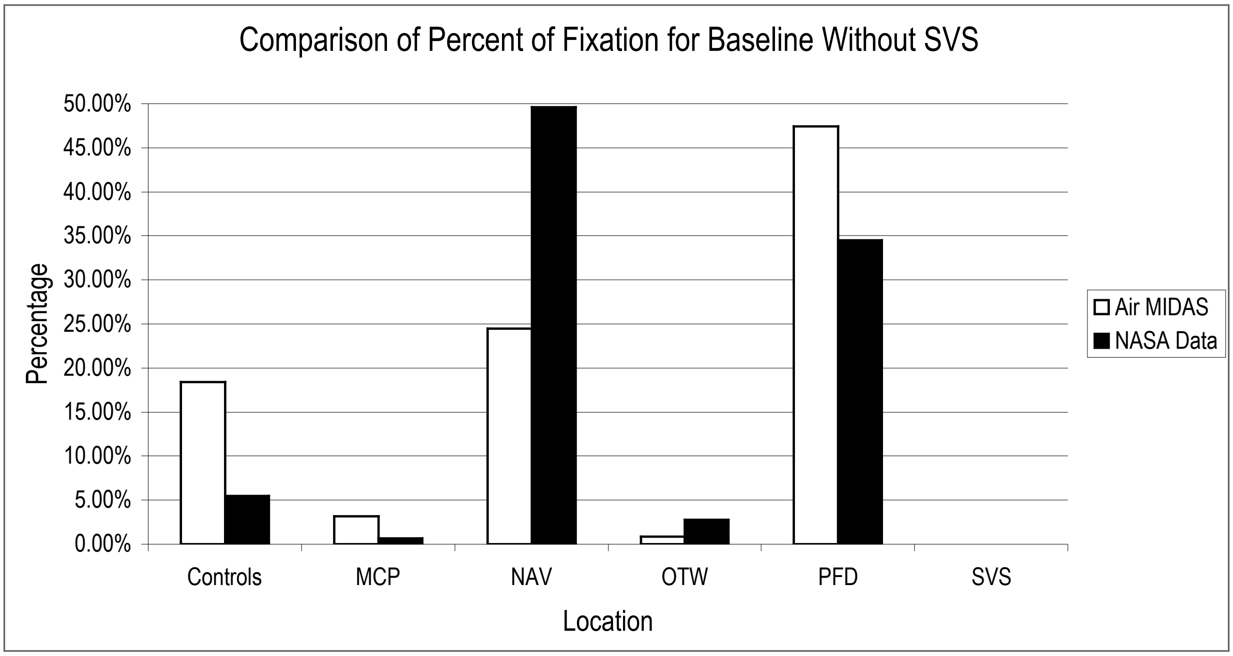


Figure 10. Model-Human Comparison of Baseline (no SVS) Fixation Percentage Location.

Figure 9 illustrate that the Air MIDAS model predicted slightly higher fixation on the controls, the MCP and the PFD than did the human data produced by Goodman, Hooey, Foyle, & Wilson (2003). The Air MIDAS model predicted lower dwells on the ND and the OTW scene than did Goodman, Hooey, Foyle, & Wilson (2003). This suggests that the rules guiding human performance are different than those guiding the model's performance. We might infer that the flight crew relies more on the information on the ND than does the Air MIDAS flight crew. Also, the Air MIDAS pilot fixated more on the PFD than does the NASA pilot (Goodman, Hooey, Foyle, & Wilson, 2003).

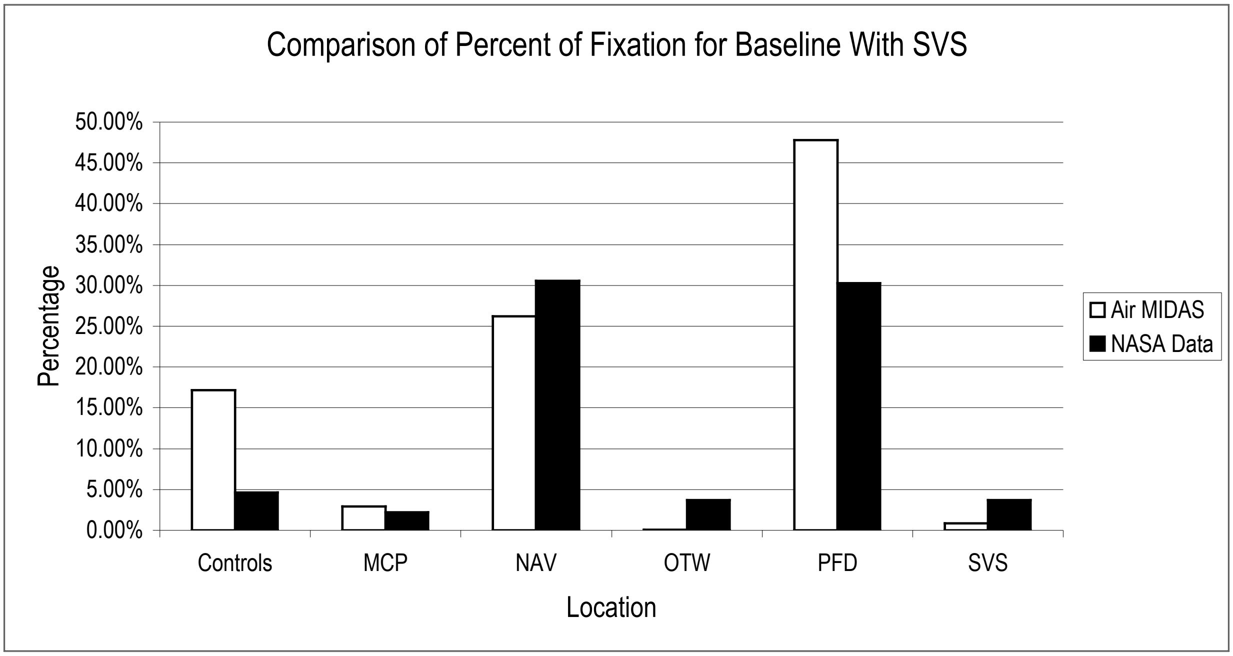


Figure 11. Model-Human Comparison of Baseline (With SVS) Fixation Percentage Location.

Figure 10 illustrates that the Air MIDAS model predicted slightly higher fixation on the controls, the MCP and the PFD than were observed by Goodman, Hooey, Foyle, & Wilson (2003). The Air MIDAS model produced lower dwells on the ND, the OTW scene and the SVS displays than did Goodman, Hooey, Foyle, & Wilson (2003).

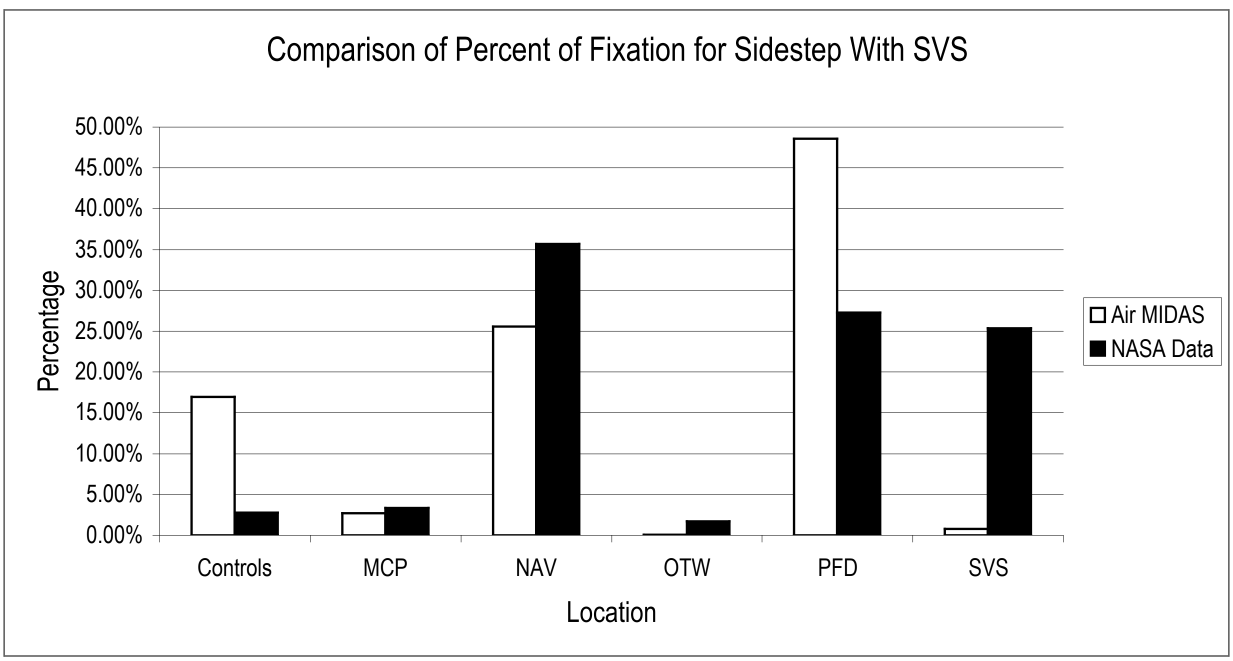


Figure 12. Model-Human Comparison of Sidestep (With SVS) Fixation Percentage Location.

The correlation of dwell time performance between human and model is lowest in the sidestep maneuver scenario as demonstrated in Figure 11. This is expected as the sidestep maneuver was least like the model baseline parameters. The kinds of information needed to support the sidestep and its implementation in SVS will need to be more closely examined in the next phase of research to better tune the model performance and dependence on the SVS system.

3.5.2 Procedural Activity Examination

Air MIDAS activities are structured in a hierarchy with goals, at the highest level of that hierarchy, being decomposed to sub-goals, and finally activities to produce the behavioral trace. We provide the goal sequence analysis to illustrate the differences in goal order between the scenarios run by the model. We compare the model-generated order to a "nominal goal ordering" based on established approach procedures.

The following figures demonstrate the order for the goals for both the pilot flying (PF) and the pilot not flying (PNF) throughout the three scenarios that were run. Figure 12 provides an outline of the predicted-operator-goal's performance in the 'Baseline without SVS condition' (Scenario 4) and provides some insight into the active goals that are being completed by the respective agents (PF or PNF) relative to the nominal baseline performance. For example, it can be seen that there is a different behavioral pattern associated with the localizer capture callout in Figure 12. The PF performs this task quickly and similar to the performance expected during the nominal condition while the PNF is unable to complete this procedure due to the unavailability of resources. This results in a longer time to completion by the simulated PNF and could highlight a potential vulnerability in system performance if procedural requirements are added to the PNF at this time. Figure 13 illustrates the model's predictions for procedural performance in the SVS operation and Figure 14 illustrates the model's predictions for the Air MIDAS operator's procedural performance in the SVS with sidestep scenario.

Nominal tasks	**Pilot Flying**	**Pilot Not Flying**
1	1	1
2	2	2
3	3	3
4	4	4
5	5	4
6	6	5
7	6	6
8	7	6
9	7	7
10	8	9
11	9	11
12	10	10
13	10	12
14	14	13

1. Set Precision FMS approach
2. 2500 callout
3. GS captured callout
4. LOC captured callout
5. Outer marker Callout
6. Final Descent Checklist
7. Receive ATC Clearance for runway
8. Check Procedure Turn
9. 1000 ft Callout
10. Set Flaps 30
11. 500 ft Callout
12. 100 ft above Decision Height Callout
13. Decision Height Callout
14. Check for stability and decide to land or go-around

Figure 13. Goal Order for Pilot Flying and Pilot Not Flying in the Baseline without SVS condition (Scenario 4).

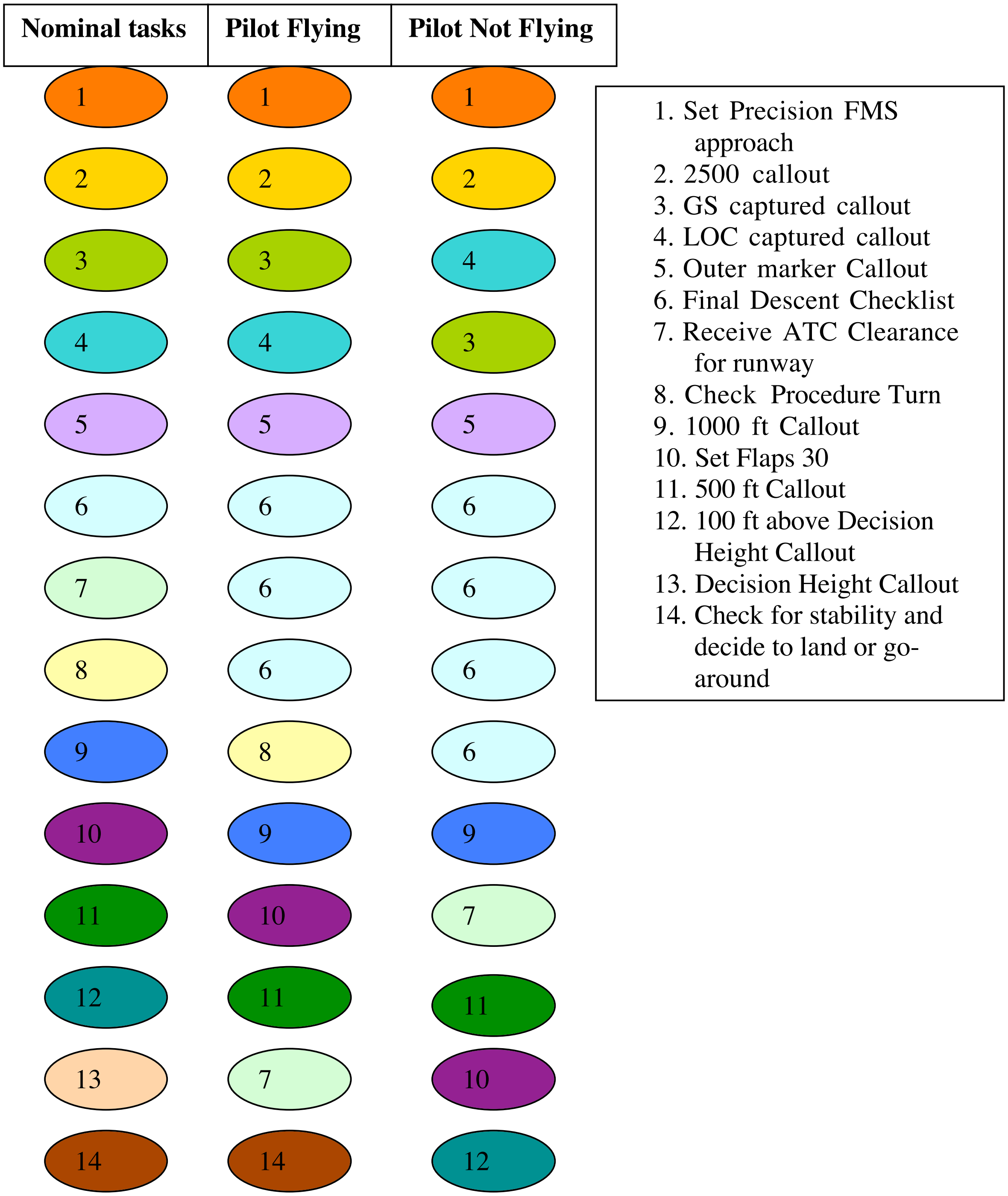


Figure 14. Goal Order for both Pilot Flying and Pilot Not flying in the Baseline with SVS condition (Scenario 7).

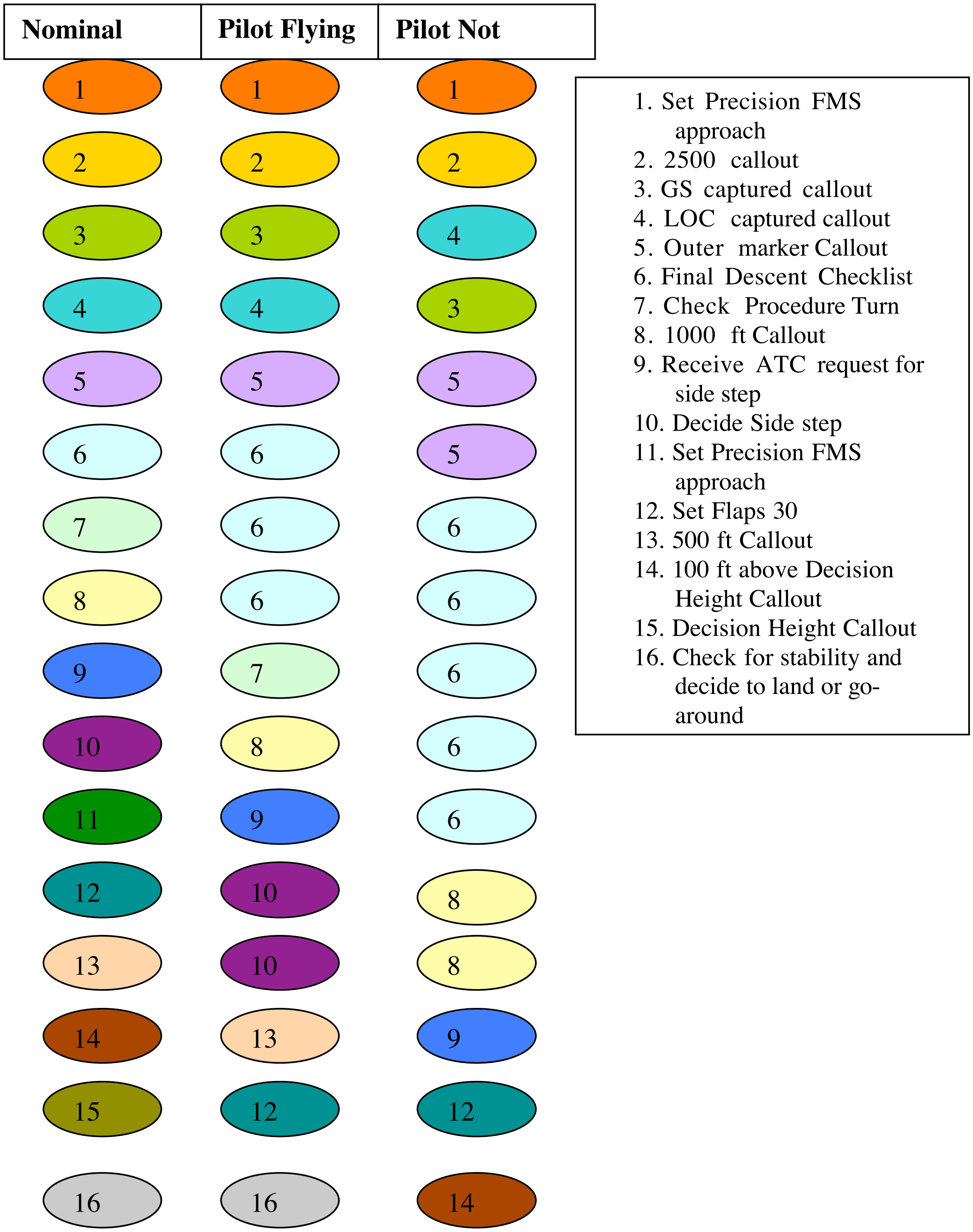


Figure 15. Goal Order for both Pilot Flying and Pilot Not Flying Sidestep Scenario with SVS (Scenario 8).

These data highlight the behavioral differences that exist between the three simulation environments programmed in the current simulation (Baseline without SVS condition, Baseline with SVS condition and the Sidestep Scenario with SVS). Behavioral changes begin to emerge when the operators are required to perform various procedural requirements in response to the environmental demands. Some of the required procedures are omitted, while others are flipped, and others still are extended. It is interesting to note that some cognitive and decision-making elements appear not to be completed. This lack of completion may lead to system vulnerabilities as the flight crew performing within the complex system appear not possess the cognitive resources to perform the activities. Goal ordering shows some evidence of early procedural completion by one of the flight crewmembers and a later completion by the other. This might suggest that one of the Air MIDAS operators may lack resources while the other Air MIDAS operator may possess sufficient resources to take over and assist the overburdened crewmember.

4 Discussion

The simulation experiment report provides support for the use of computational human performance models in system design and analysis. The validation effort provides evidence that the Air MIDAS tool with its constituent models of vision, audition, perception, attention, and its cognitive architecture generates behavior that is similar to human-in-the-loop performance. The performance differences that emerge in the current simulation provide insight into the simulation processes that could benefit from further work.

4.1 Simulation Data Generation Speed

The complexity of the operating environment and the level of detail required to update the worlds of the agents in the simulation resulted in slow computational performance in the data generation. We will explore methods to produce a more computationally efficient program.

4.2 Air MIDAS Model Development

There was a significant challenge involved in synchronizing the Air MIDAS equipment data with aircraft state/equipment data obtained from the NASA part task simulation. Goodman, Hooey, Foyle, & Wilson (2003) collected the simulation data every 10 ms, whereas the tick resolution used by Air MIDAS is 100 ms. As a result, there was effort involved in data reduction and data management to synchronize the part task simulation data with Air MIDAS model data.

The initial representation of the information accuracy function was the same for both the "non-directed" visual sampling, i.e. general scan, and for the "directed" information-seeking behavior. This rule was modified in our simulation runs to enable goal-directed behavior to always perceive information accurately.

4.3 Future Research Considerations

Visual target detection was noted as being a difficult task to incorporate into the human performance model. Landy (2002) provided equations to incorporate a model of target detection. Inclusion of this model would be a benefit for the modeling software and will be explored in future modeling efforts.

It also became apparent in working through the requirements to incorporate vision into a human performance model that representing human depth and distance is a significant challenge that will need to be addressed in the next phase of research in human-system simulation.

5 References

Atencio, A. (1994). Air Warrior MIDAS. [On-line]. Available. *http://caffeine.arc.nasa.gov/midas/original-midas/Air_Warrior.html.*

Atencio, A. (1998). Short Haul Civil Tiltrotor MIDAS. [On-line]. Available. *http://caffeine.arc.nasa.gov/midas/original-midas/Tiltrotor_MIDAS.html.*

Arditi, A. & Azueta, S. (1992). Visualization of 2-D and 3-D aspects of human binocular vision. Presented at the Society for Information Display International Symposium. The Lighthouse, Inc.

Balci, O. (1998). Verification, validation, and testing techniques. In J. Banks (ed.) Handbook of simulation: Principles, methodology, advances, applications, and practice, pp. (335-427), N.Y.: Wiley and Sons, Inc.

Corker, K. (2000). Cognitive Models & Control: Human & System Dynamics in Advanced Airspace Operations. In N. Sarter and R. Amalberti (Eds.) Cognitive Engineering in the Aviation Domain. Lawrence Earlbaum Associates, New Jersey.

Corker, K. (1994). Man-machine integration design and analysis system (MIDAS) applied to a computer-based procedure-aiding system. In *Proceeding of the 38th Annual Meeting of the Human Factors and Ergonomics Society.*□Human Factors and Ergonomics Society, Santa Monica, CA, October, 195-198.

Corker, K.M., Gore, B.F., Fleming, K. & Lane, J. (2000). Free flight and the context of control: Experiments and modeling to determine the impact of distributed air-ground air traffic management on safety and procedures. In *Third Annual Eurocontrol International Symposium on Air Traffic Management*, Naples, Italy, June 13-16.

Corker, K.M. & Guneratne, E. (2002). Human Factors Issues and Evaluation for Commercial and Business Aircraft Synthetic Vision Systems. HAIL Laboratory Report #SVS-2-2002. San Jose State University, San Jose Ca.

Gore, B.F. (2002). Human performance cognitive-behavioral modeling: A benefit for occupational safety. In B. Chase & W. Karwowski (Eds.), International Journal of Occupational Safety and Ergonomics (JOSE), 8 (3), 339-351.

Goodman, A., Hooey, B.L., Foyle, D.C., & Wilson, J.R. (2003). HPM-SVS part-task simulation documentation: Characterizing pilot performance during approach and landing with and without visual aiding. Moffett Field, CA: NASA Ames Research Center.

Landy, M.S. (2002). Vision and attention for Air Midas (NASA Final Report NCC2-5472). Moffett Field, CA: New York University, Dept. of Psychology and Center for Neural Science.

Law, A.M., & Kelton, D.W. (2000). Simulation, modeling and analysis (3rd Edition), N.Y.: McGraw Hill.

Lubin, J. & Bergen, J. (1992). NASA cockpit display visibility modeling project (NASA Final Report NAS2-12852). Moffett Field, CA: SRI/David Sarnoff Research Center.

McCracken, J.H. & Aldrich, T.B. (1984). Analysis of selected LHX mission functions: Implications for operator workload and system automation goals (Technical note ASI 479-024-84(b)). Anacapa Sciences, Inc..

Mumaw, R., Sarter, N., Wickens, C., Kimball, S., Nikolic, M., Marsh, R., Xu, W., & Xu, X. (2000). Analysis of pilot monitoring and performance on highly automated flight decks (NASA Final Project Report: NAS2-99074). Moffett Field, CA: NASA Ames Research Center.

National Academy Press (1990). Quantitative Modeling of Human Performance in Complex, Dynamic Systems. In S. Baron, D.S. Kruser, & B.M. Huey (Eds.), Panel on Human Performance Modeling, Committee on Human Factors, and the Commission on Behavioral and Social Sciences and Education, National Research Council, Washington, D.C.

Remington, R., Johnston, J. & Yantis (1993). Involuntary attentional capture by abrupt onset, Perception and Psychophysics, 51(3), 279-290.

Modeling the NASA Baseline and SVS-Equipped Approach and Landing Scenarios in D-OMAR

Stephen Deutsch and Richard Pew
BBN Technologies
Cambridge, Massachusetts

1. Introduction

Human performance modeling (HPM) is a technology that has the potential to help address flight deck and air traffic controller (ATC) workplace design issues and support aircrew procedure evaluation. It can be used early in the workplace development process before prototypes or full-scale simulations are available or later in the design cycle to support on-going development. Unfortunately, the technology is not yet sufficiently mature that it can be applied easily and routinely. However, it is reasonable to undertake exploratory studies to evaluate aircrew procedures for employing new flight deck technologies and to evaluate their role in promoting aviation safety. At the same time, these investigations will improve the architectures and software tools that are available to support the human performance modeling process.

Under the aegis of the Aviation Safety Program, NASA Ames Research Center (ARC) has initiated a program element concerned with modeling human performance. The element's goals include advancing the state-of-the-art in human performance modeling and demonstrating their potential payoffs in the commercial aircraft design and aircrew procedure development by contributing to improved aviation safety. Within the framework of modeling human performance, achieving a better understanding of the sources of human error and identifying procedures for error mitigation are a particular focus. This element is a multi-year, multi-contractor effort that is just completing its second year.

For the second year effort, NASA asked that the modeling teams examine approach and landing operations, and compare aircrew performance for baseline and synthetic vision system (SVS) equipped flight decks. BBN's effort has focused on the further development of captain and first officer human performance models and the approach and landing procedures that employ baseline and SVS-equipped flight deck configurations. Human performance models for the approach, tower, and ground controllers have been developed to interact with the aircrew models to faithfully represent the airspace. The human performance modeling effort paralleled and profited from the NASA part-task simulation studies (Goodman, Foyle, Hooey, & Wilson, 2003) that collected pilot performance data in baseline and SVS-equipped flight deck operations at the Santa Barbara Municipal Airport (SBA).

The modeling effort was accomplished using BBN's Distributed Operator Model Architecture[1] (D-OMAR) to represent the behaviors of the aircrews and air traffic controllers. D-OMAR was also used to model the aircraft and their flight decks, the ATC workplaces, and the essential features of the Santa Barbara Municipal Airport and the local airspace. Our goal for this year was to produce models that could appropriately represent successful approach and landing performance similar to that exhibited by the pilot subjects in a subset of the recently conducted NASA SVS part-task simulation experiments. To date, the D-OMAR models are successfully executing five of the part-task scenarios.

In this document, Section 2 traces the development of the human performance models, the aircrew and air traffic controller procedures that they execute, and the strategy for building the emulation of the part-task scenarios. Section 3 then provides a discussion of the findings that resulted from the modeling effort, Section 4 outlines the implications of the findings, and

[1] The D-OMAR software with user manual is available as OpenSource at http://omar.bbn.com.

Section 5 highlights the lessons learned in the process of developing the human performance models to implement the scenarios.

2. Modeling the Approach and Landing Scenarios

2.1 NASA Support for the Modeling Effort

NASA Ames made available several important resources to facilitate this year's modeling effort. The part-task simulation dictated an RNAV rather than an ILS approach as modeled for last year's O'Hare scenarios (Deutsch & Pew, 2001: Deutsch & Pew, 2002). To assist the modeling teams in the transition, a cognitive task analysis (Keller & Leiden, 2002a) provided a detailed description of aircrew and air traffic controller procedures for an RNAV approach. The document included detailed information on how flight deck systems are used to support an RNAV approach. An addendum to the document (Keller & Leiden, 2002b) extended the task analysis to cover the aircrew's use of an SVS during the approach and landing. These documents provided much of the information necessary to construct the D-OMAR goal, sub-goal, and procedure framework that represents the standard aircrew and air traffic controller procedures for an RNAV approach and landing.

Our modeling effort relied heavily on the documentation of and data from the NASA baseline and SVS part-task scenario trials. Goodman *et al.* (2003) provided a detailed description of the simulated flight deck, the design for the ten scenario trials, and a description of the scenario trial data for three subjects. Trial data included simulation output for each run, eye tracker data, and video (with audio) recordings based on an eye-tracker camera and a room-view camera. The eye tracker data included both fixation sequence and dwell sequence data files. The data was made available by trial with summary statistics available by phase of flight. A summary fixation sequence spreadsheet was assembled to provide duration percentage for viewing each instrument across all trials and all subjects. The spreadsheet made it possible to compare and contrast individual subject eye tracking behaviors across the trials. In summary, these data provided a comprehensive view of aircrew behaviors essential to modeling the baseline and SVS-assisted approach and landing trials.

The BBN team also participated in and profited from an SVS/SWAP information-sharing workshop held at NASA Langley late in 2001. The workshop provided a number of presentations on the development of the NASA SVS system as well as a review of the flight test experiments in using the system. The SVS section of the NASA Aviation Safety Program (AvSP) web pages provided access to a number of publications that included the Concept of Operations for Commercial and Business Aircraft Synthetic Vision Systems (Williams, Waller, Koelling, Burdette, Doyle, Capron, Barry, & Gifford, 2001).

2.2 Strategy for Developing the Scenarios

The HPM-SVS part-task simulation study (Goodman *et al.*, 2003) included ten scenarios that selectively covered three variables of interest: display configuration, visibility, and approach events. There were two display configurations: a baseline flight deck consisting mainly of a primary flight display (PFD), a horizontal situation indicator (HSI), and a mode control panel (MCP), and a second configuration in which the baseline configuration was augmented with an SVS display. Landing gear, flaps, speed brakes, throttle settings, and map scale for the HSI were set by the first officer using a separate display panel.

Visibility conditions included visual meteorological condition (VMC) and instrument meteorological condition (IMC). The light haze in VMC allowed visual flight rules. Under IMC, there was a *reported* 800-foot ceiling. Approach events included a nominal approach, a "late reassignment" requiring a sidestep to a parallel runway, a missed approach requiring a go-around, and a "terrain mismatch" in which the SVS was found to be misaligned as the aircraft emerged from the cloud cover. The "go-around" scenarios took two forms: in the VMC case, traffic on the runway made the go-around necessary. In the IMC case, in spite of the reported 800-foot ceiling, the aircraft was still in the cloud cover at the 650-foot decision height, hence the captain was unable to acquire the runway making the go-around necessary.

As modelers, we were asked to focus on the "late reassignment" scenarios using the baseline and SVS-equipped flight decks. Our approach was to first develop the basic aircrew procedures to support the RNAV approach and landing using the baseline flight deck under VMC. We then added weather to represent IMC. In the modeled IMC condition, there was an *actual* cloud ceiling at 800 feet; hence, the breakout from the cloud cover occurred at 150 feet above the designated decision height.

For the "late reassignment" condition, we developed a scenario with two closely spaced aircraft on the approach to SBA 33L. As the lead aircraft landed, it blew a tire and temporarily held on the active runway. This set up a situation requiring the ATC to request that the following aircraft side step to the parallel runway 33R. Lastly, we added an SVS to the flight deck and extended the aircrew procedures to include the use of the SVS for the IMC scenario and the "late runway reassignment" scenario.

At this point, the modeled aircrews successfully execute the approach, landing, and taxi procedures for five of the NASA part-task simulation scenarios:

- Scenario 1 – nominal approach using the baseline configuration in VMC
- Scenario 2 – late reassignment approach using the baseline configuration in VMC
- Scenario 4 – nominal approach using the baseline configuration in IMC
- Scenario 7 – nominal approach using the SVS configuration in IMC
- Scenario 8 – late reassignment approach using the SVS configuration in IMC

2.3 Human Performance Models for the Aircrew and the Air Traffic Controllers

In the scenarios, the captain and first officer must work together to safely execute the approach-and-landing procedures. In responding to a series of controller directives, actions must be prioritized, appropriate aircrew communications must be generated to coordinate the execution of these actions, and interrupts in the form of further directions from the controllers must be handled. The interrupts are not unexpected, but rather meet expectations consistent with the local air traffic and weather conditions. Reactive behaviors are determined within the framework of the aircrew's active goals and procedures. In meeting their responsibilities, the captain and first officer have a significant number of tasks in process, each of which requires a coordinated mix of perceptual, cognitive, and motor skills. The scenarios create situations in which the response to demands must be carefully prioritized to achieve acceptable performance.

Each aircraft is populated by human performance models for the captain and first officer. The aircrew models are extensions of last year's models that executed the O'Hare ILS approach, landing, and taxi (Deutsch & Pew, 2001: Deutsch & Pew, 2002). The new RNAV procedures that they employ at SBA are based on the Keller and Leiden (2002a; 2002b) cognitive task analysis. Consistent with our long-term goal of examining error mitigation for two person aircrews, the procedures that our models execute closely follow the Keller and Leiden real world task analysis. Hence, there are departures from the exact procedures as tailored for the part-task simulation trials where the captain was the experiment subject and the first officer was a surrogate minimally supporting the captain as necessary.

In the same spirit, the modeled approach and landing scenarios follow the standard progression in the airspace transiting from approach controller to tower controller to ground controller and terminating as the aircraft completes the landing and taxies to its assigned concourse. In the modeled scenarios, each controller is represented by a human performance model. The scenarios as modeled more closely follow real world operations than was possible in the part-task simulation.

The aircraft model includes the instruments and controls necessary for the crew to execute the required approach and landing scenarios. The model more closely emulates the actual 757 flight deck rather than part-task simulation flight deck. The principal instruments include the PFD, the HSI, and the SVS. The PFD includes annunciators for LNAV and VNAV. Controls include the MCP, switches for the autopilot, and levers for the throttles, flaps, landing gear, and speed brakes. The central instrument panel includes lights providing landing gear status.

The aircrew makes use of the approach plate for SBA runway 33L for information related to the RNAV approach. The information used by the captain to brief the go-around contingency at the beginning of the approach is derived from the approach plate. The approach plate is subsequently used by the aircrew as a reference to track the sequence of fixes for the approach path and as the source for the required descent altitude for each approach leg. As they depart the active runway, the first officer uses the SBA airport diagram to support taxi operations leading to the concourse.

Voice communication between the captain and first officer is used to coordinate the execution of approach and landing procedures. Party-line radio communication is modeled with the aircrew resetting radio frequencies as they move from one controller to the next. Careful attention has been paid to the fine details of interleaving of aircrew and air traffic controller conversations and to handling air traffic controller interruptions to aircrew conversations.

As each scenario begins, the approach controller clears NASA186 (and NASA277 for the two aircraft scenarios) for the approach to SBA 33L and provides information on VMC or IMC depending on the scenario. Using information from the approach controller and the approach plate for runway 33L, the captain continues the approach by reviewing the runway and weather information with the first officer. Go-around procedures are briefed, based on the approach plate information. For the RNAV approach, the captain asks that MCP modes LNAV and VNAV be set and checks the PFD annunciator mode lights.

The aircrew then focuses on navigation as the aircraft proceeds along the flight management computer (FMC) flight path from one fix to the next. As they approach each fix, the captain calls for a new MCP altitude setting for the next leg. The aircrew monitors the aircraft's heading and altitude changes (information derived principally from the HSI) as the aircraft transitions onto the leg to the next fix. They continue to monitor the heading until the new desired heading is fully established. They monitor altitude to assure that they hold at the designated target altitude. As the approach progresses, the captain calls for a series of flap settings consistent with their speed and position along the approach path. The handoff from the approach controller to the tower takes place along the leg to GOLET.

The aircrew contacts the tower and is immediately given the clearance to land. The captain asks for the final flap setting, that the landing gear be lowered, and that the speed brakes be armed. At this point, the captain asks for execution of the landing checklist, which is then acted on with the first officer. As the aircraft's descent continues, the first officer monitors the aircraft's altitude and makes call outs at 1000 feet above field level (afl), as they approach decision height, and at 100 feet afl. The captain is responsible for the out-the-window sighting of the runway and making the decision to land. Under VMC, the captain can readily acquire the runway out the window well before decision height. For the SVS-equipped, IMC-condition scenario, the captain can use the SVS to acquire the runway before they break out of the cloud cover, but must still acquire the runway out the window to make the decision to land. The captain anticipates the break out from the cloud cover at 800 feet and has adequate time to acquire the runway before the 650-foot decision height. The captain must take manual control of the aircraft to preempt the preprogrammed go-around and further manage the landing.

Figure 1 provides a D-OMAR plan view of two aircraft on their approach to SBA 33L in the "late reassignment" scenario. At the point presented in the plan view, the first aircraft has blown a tire on landing causing it to hold temporarily on the active runway. The top panel on the right in Figure 1 records the conversations among the approach and tower controllers and the two aircraft. NASA277's communication with the tower controller related to the blown tire shows up in the recorded dialog. The tower controller addresses the situation by asking the second aircraft, NASA186, to sidestep to runway 33R.

The lower panel on the right records details of the conversation on the flight deck of the following aircraft, NASA186. Diane, the NASA186 captain, tells her first officer that she will accept the request to sidestep to 33R. The first officer communicates the captain's acceptance of the sidestep request to the tower controller. The first officer's response completes the sidestep transaction with the tower controller related in the upper display panel. The captain then announces her decision to proceed with the landing. As they land, the captain manages speed brake and reverse thrust settings, while the first officer calls out current ground speed. Following the landing, NASA186 notifies the tower as the aircraft leaves the active runway, contacts ground controller and receives directions to taxi to the it's assigned concourse.

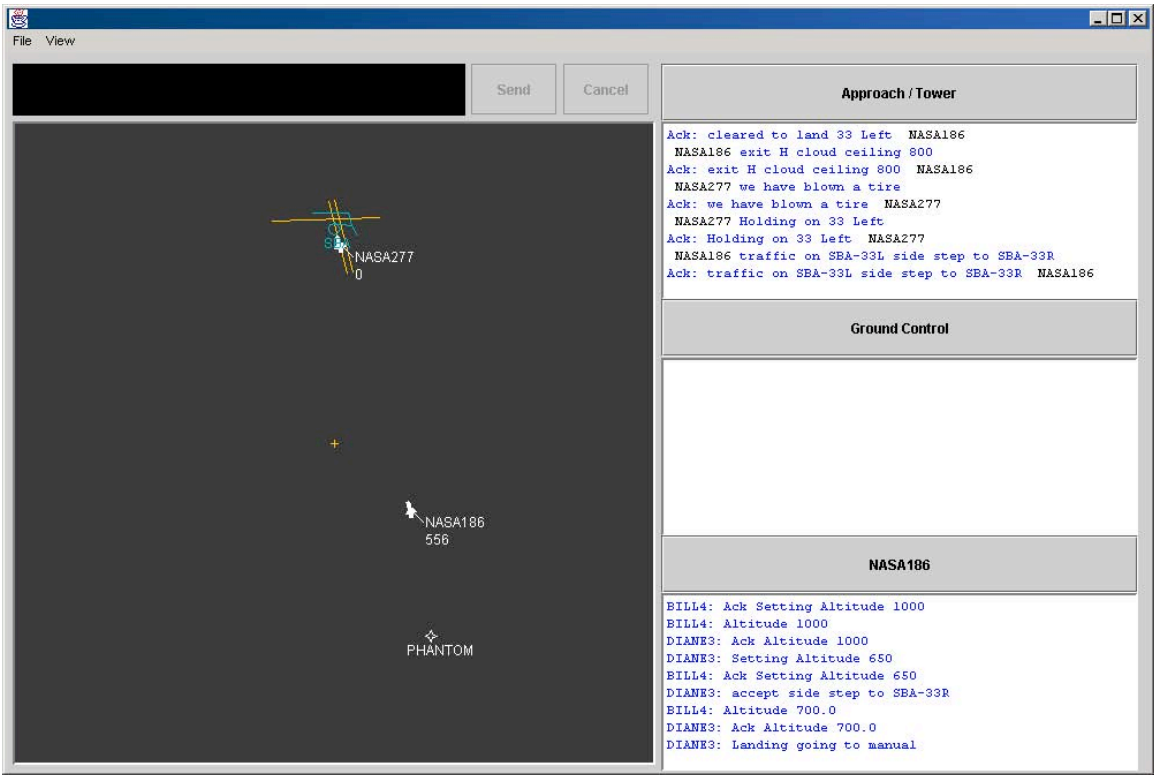


Figure 1. Screen View from the Late Reassignment Scenario.

2.4 Assessing Model Behaviors in D-OMAR

D-OMAR simulation tools provide explicit measures of model behaviors. They are essential in the assessment of model performance just as they are essential for managing the complexity required to create the models. Aircrew behaviors are frequently multitask behaviors—the more, or occasionally less, successful integration of the demands of several ongoing procedures. Each of these procedures is made up of several steps that require the coordinated interaction of several human functional capabilities (e.g., the maintenance of a conversation, the coordination of hand-eye actions to set a selector). The execution of a checklist interrupted by an ATC communication is at once a common occurrence and a challenging event sequence to faithfully model. Multiple levels of visibility into model behaviors are essential both to develop the scenarios and to assess model performance.

D-OMAR graphical display tools each provide a unique view into model behaviors. A plan view, as illustrated in Figure 1, allows an observer to monitor the progress of the aircraft along its flight path. The plan view display has recently been supplemented by a similar HSI-like display (see Figure 4). A Gantt chart display (see Figure 3) provides detailed information on goals and procedures as executed by the captain and first officer. An event timeline (not illustrated) provides detailed insight into the behaviors of the publish-subscribe protocol used to coordinate procedure execution. Lastly, a detailed event trace is recorded for each simulation run with key events displayed in the trace pane of the simulation control panel (see Figure 2) as the simulation progresses.

Some of the evaluation tools operate concurrently with the simulation; others are used once a simulation run has completed or after the simulation is paused. The plan views and the simulation trace operate concurrently with the simulation, the task and event timelines are available once the simulation has been paused. An event recording system is used to capture the data to support most of the evaluation tool presentations. Several event types are basic elements of the D-OMAR simulator, others are more specialized and created to address the data capture requirements of a particular domain or scenario. Procedure execution is the basic element driving agent behaviors; hence, events are recorded that identify the agent executing the procedure, the beginning and end times for each procedure, the success or failure of the procedure, the procedure's priority, the name of its parent procedure, and any time periods during which the procedure was interrupted.

Several event types have been created to track the performance of the human performance models for the aircrew and air traffic controllers. One of the event types records flight deck actions taken by each aircrew member (and workplace actions for the air traffic controllers); for example, the setting of an MCP selector for altitude or the movement of the lever to establish a particular flap setting. Since in-person conversations on the flight deck and party-line conversations with the air traffic controllers are so important, conversation events have been defined to record these conversations. The aircrew and air traffic controller conversations presented in the right hand panels of Figure 1 are generated by "after methods" on the event recording process.

The event types for flight deck actions and conversations are key elements providing data for the on-line trace. They record and print the actions taken by the aircrew and the air traffic controllers. As such, they represent the outcomes of the execution of the goals and procedures whose development was based on the cognitive task analysis for the RNAV approach and landing (Keller & Leiden, 2002a).

The simulation control panel (Figure 2) provides an interface for the user to select and manage the execution of a scenario, and an on-line trace of selected scenario events. The "Scenario" line in the panel provides for the selection of the scenario to be executed. The "Initialize," "Run," and "Pause" buttons enable the user to control the flow of scenario execution. The D-OMAR simulator is capable of both real-time and fast-time operation. In fast-time mode, the simulator is very efficient taking about 40 seconds to complete 1150 seconds of real-time for the two aircraft "late reassignment" scenarios.

The sample panel shown in Figure 2 includes a short section of the trace from the nominal VMC approach scenario, the SBA-RNAV-VMC-SCENARIO. At this point in the scenario, the captain has just asked that the landing gear be lowered, the flaps be set to 25 degrees, and the speed brakes be armed. As recorded in the trace, the first officer attends to each request and executes each requested action, first checking current setting and then adjusting the setting as necessary. Part way through the sequence of requests for actions on the part of the first officer, the captain adjusts the MCP IAS/mach speed selector and consistent with established procedures, announces the change as it is made.

The slider at the side of the Figure 2 provides a hint at the size of the trace for the nominal VMC scenario. The trace is lengthy, but manageable in size, in large part, because the printing for most event types is turned off. For example, procedure event data that is essential to the task timeline described below, would if printed, make it far more difficult to isolate the important events currently presented. The ability to capture a complete set of scenario event data, yet tailor the trace content is important in making the trace a useful analysis tool.

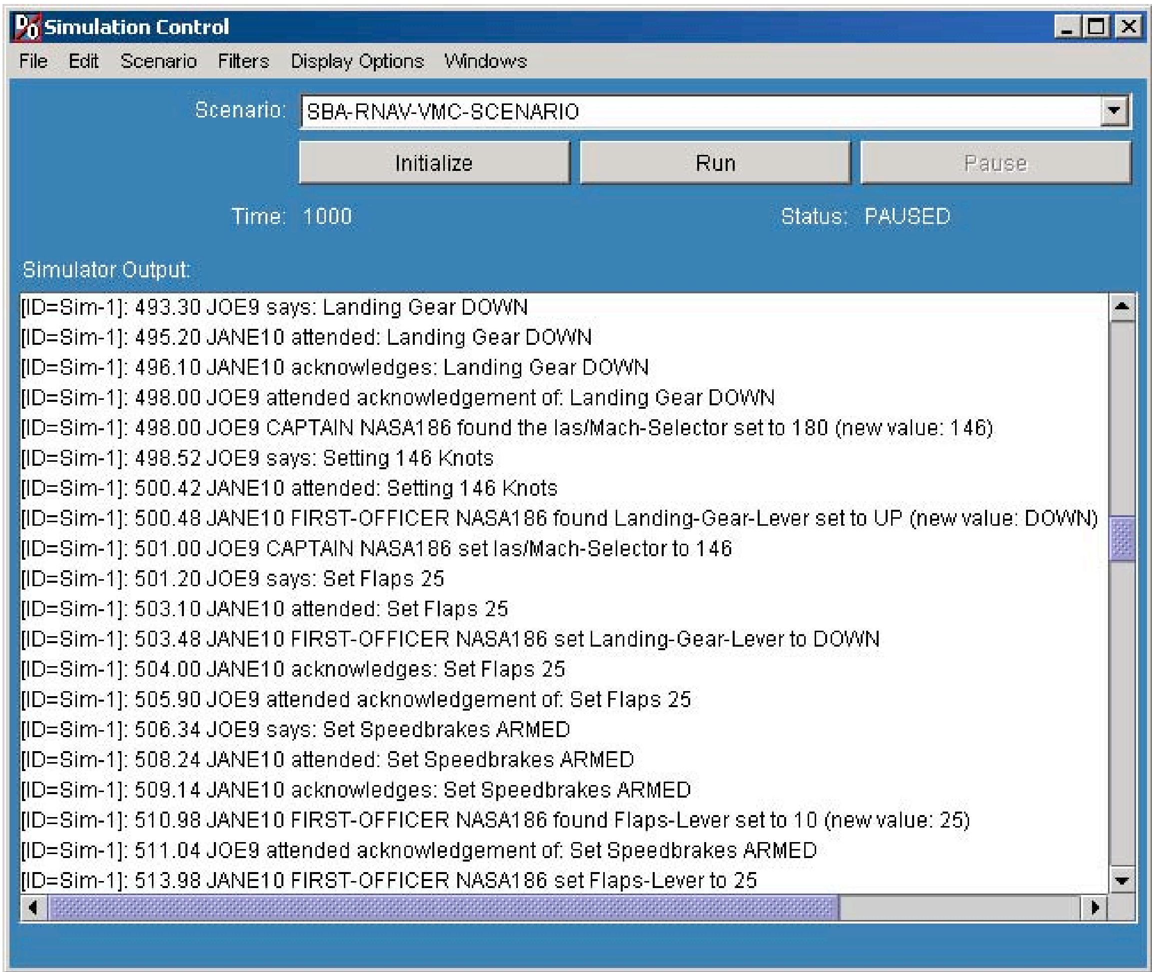


Figure 2. Simulation Control Panel and On-line Trace.

The on-line trace has been specialized to provide insight into critical actions taken by aircrew and air traffic controller models in the scenarios. The actions are the products of the execution of clusters of a captain's or first officer's goals and procedures that each has duration in time. The task timeline display, a Gantt-style display is used to provide insight into how an agent's goals and procedures play out in time to generate these actions. Figure 3 is an example of the timeline display from the baseline VMC scenario. The slider at the right of the figure is quite small indicating that the captain has quite a large set of on-going goals and procedures beyond those currently visible in the display. Some of these goals and procedures represent preparedness to respond to anticipated events, some are currently active in the timeframe covered by the display. The selected timeframe for the display, identified in the display's bottom panel, is fifty seconds, from 550 seconds into the scenario to 600 seconds in the scenario. The aircraft is approaching GOLET and about to transition to the leg to PHANTOM on the approach to SBA 33L.

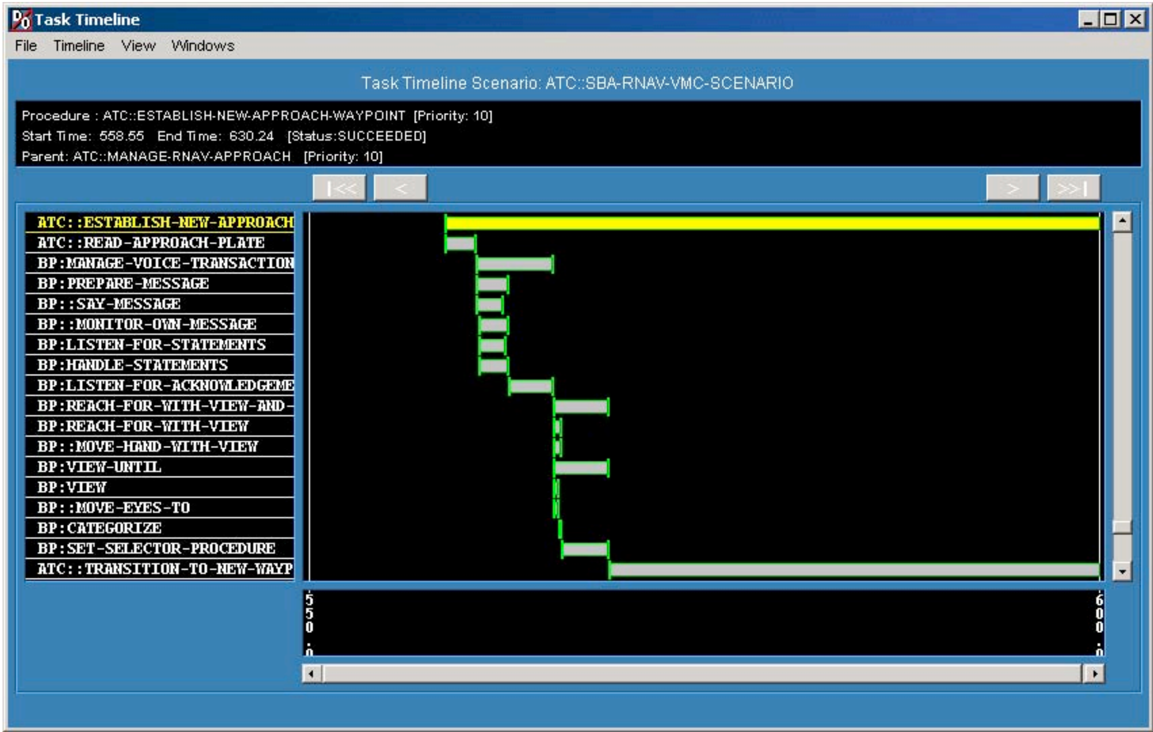


Figure 3. Task Timeline Display.

The procedures that appear in the timeline in Figure 3 are a subset of the captain's procedures as s/he prepares for and then monitors the transition to PHANTOM, the next leg of the approach. Each line of the display represents one of the captain's goals or procedures. The goal or procedure's name appears in the panel at the left, the bar represents the duration of the procedure's execution. Time periods for which a procedure is interrupted (there were none for these procedures) are indicated within the bar representing the procedure's duration. The mouse (not shown), over the first procedure, causes the top panel to be filled with that procedure's primary attributes: the procedure's name and priority, its start time, its end time if it has completed, its current status, and the name of its parent goal or procedure.

In the procedures shown in Figure 3, the captain is concerned with the basic actions for establishing the aircraft on the new leg. The captain first reads the approach plate to verify the altitude for the next waypoint (executed as a single procedure). The captain then tells the first officer that s/he is setting the MCP with the new desired altitude (the next seven lines of the display cover the verbal transaction: speaking the message and listening to the acknowledgement). The captain then sets the MCP IAS/mach speed selector (a coordinated hand-eye activity accomplished by the next eight lines of the display). In the last line shown in the timeline display, the captain is concerned with monitoring the aircraft's transition to the new heading and altitude. Information on current heading is derived from the on-going scan of the HSI display. Information on current altitude is derived from the PFD and the SVS for the scenarios in which it is available. The monitoring procedure obtains this information by "subscribing" to scan procedures that "publish" this information. The publish-subscribe protocol provides a basic mechanism to coordinate procedure execution and to move information between procedures. It is more fully described in the following section.

2.5 Describing the D-OMAR Human Performance Models

The development of human performance models in D-OMAR (Deutsch & Pew, 2002; Deutsch & Pew, 2001; Deutsch, 1998; Deutsch & Adams, 1995) has been based on research in cognitive neuroscience, cognitive science, experimental psychology, and recent cross-disciplinary work in the theory of consciousness. As with most human performance models, the complexity of the models makes it difficult to provide a description that is both brief and complete. There is a theoretical framework that underlies the architecture, a broad range of individual human functional capabilities that must be represented, there are complex interactions among these capabilities, that taken together, generate the observed behaviors, and there are practical compromises that inevitably must be made in producing a working model. To provide additional insight into the D-OMAR human performance model behaviors used in the NASA SBA approach and landing scenarios, it might be useful to briefly examine a few central features of the model: how multiple task behaviors are modeled, the role of vision in supporting the model's multiple task behaviors, and the modeling of working memory.

2.5.1 An Aside on Model Architecture

For convenience within this report, we have spoken of the D-OMAR model, but to be more accurate, that reference needs clarification. D-OMAR itself is simply a general-purpose discrete event simulator. It has been tailored specifically to provide a software framework in which to explore alternate architectures for human performance modeling. The D-OMAR representation languages, a frame language, a rule language, and a procedural language[2] provide the basis for constructing the alternate architectures. The particular models employed for this NASA research task are a further development within an architecture for human performance modeling that has evolved over a number of years.

Most human performance models (e.g., ACT-R (Anderson & Lebiere, 1998), SOAR (Laird, Newell, & Rosenbaum, 1987), EPIC (Meyer & Kieras, 1997), MIDAS (Corker & Smith, 1993)) are implementations of a particular cognitive architecture. D-OMAR, rather than being a particular cognitive architecture, is a simulation framework in which to experiment with and evolve architectures for human performance models. It has been used, in this case, to implement a particular architecture that has evolved to address the NASA task. D-OMAR has readily been used to explore variations on elements of the current architecture and to implement an entirely different architecture. This level of flexibility in model architecture seems essential to the effort to improve the capabilities of human performance models.

2.5.2 Multiple Task Management

One of the principle areas of research in the development of D-OMAR has been in the area of modeling human multitask behaviors. In developing D-OMAR, we have sought to provide a computational framework in which to assemble functional capabilities that operate in parallel, subject to appropriate constraints, and that taken together exhibit the multiple task behaviors of human operators—aircrews and air traffic controllers. The desired behaviors have a combination of proactive and reactive components. That is, the operators have an agenda that they are pursuing, but must also respond to events as they occur. Consequently, within the proactive agenda, there may be newly motivated tasks for which on-going tasks must be deferred. The bounds on what can be accomplished concurrently take several forms. A

[2] Detailed information about the representation languages is available in the documentation for the Lisp version of D-OMAR at the web site http://omar.bbn.com.

typical behavior may be to set aside a flight deck conversation in order to respond to an ATC communication, while at another level, two competing tasks may each require the eyes to guide a manual operation. In the first instance, it is a matter of protocol, in the second, contention for a physical resource.

The core of a D-OMAR model is a network of procedures whose signal-driven activation varies in response to events that are proactively channeled to achieve the operator's goals. From a bottom up perspective, there is an assembly of individual perceptual, cognitive, and motor capabilities that are recruited as procedures to address current goals and sub-goals. Neumann's (1987) functional view of attention, and the localization of mental operations in the brain, as put forward by Posner, Petersen, Fox, and Raichle (1988) are important contributions supporting this approach to modeling human behaviors. Taken together, they point to the functional components in task execution as taking place at particular local brain centers with the coordinated operation of several such centers being required to accomplish any given task. The form that the coordination might take is of particular importance in developing a model of behaviors. A publish-subscribe protocol provides the signal-driven activation needed to coordinate the actions of the various perceptual, cognitive, and motor centers acting in support of the completion of the task. The publish-subscribe protocol also serves to move information among the functional centers.

From a top down perspective, the things that person knows how to do, basic person skills (e.g., coordinated hand-eye actions to set a selector) and domain specific skills (e.g., making the decision to land), are represented as goals, sub-goals, and procedures. Active goals represent the operator's proactive agenda for managing his or her tasks. The goals typically activate a series of sub-goals and procedures. The goals and sub-goals represent the objectives of the actions to be taken; the procedures are the implementation of the actions to achieve the goals and sub-goals. The procedures each may include decision points to address variations in the local situation. Hence, the operator's overall agenda is implemented by the network of procedures established by the goal hierarchy and linked by the publish-subscribe protocol. A subset of the procedures are active, most are in a wait-state—they represent the potential downstream actions of the operator's current actions and his or her ability to cope with a changing world.

Within this framework, process (Edelman, 1987; 1989) has a preeminent role. Basic person skills and domain specific skills encompass far more than simple perceptual or motor skills, they include the highly refined *cognitive* skills that are the mark of significant human expertise (Logan, 1988; Bargh & Chartrand, 1999). Taken together, a model's goals and procedures, the capabilities of the model to perform in a human-like manner, are a major component of the model's long-term memory.

2.5.3 Vision as a Component within Multi-task Behaviors

Demands on the pilot's visual system are varied and complex. A broad range of these visual capabilities is included in the models. Some of a pilot's actions are purely visual. There are basic processes that take in information from the major flight deck instruments and the view out the windscreen. The viewing of a flight deck instrument can be a generic guidance or navigation status check, or the monitoring of an instrument for a specific target value. The view out the windscreen can be to assure that there are no traffic conflicts (i.e., seeing *nothing* can be the desired outcome), acquiring a specific object (e.g., sighting runway 33L on the approach to support the decision to land), or tracking the aircraft's alignment with runway

33R while executing the sidestep maneuver. Acquiring a specific value occurs at brief instant in time, in tracking the runway, the viewing may extend over a fair period of time with breaks to address other visual tasks.

For some actions, the visual component plays a supporting role. The execution of flight deck operations require coordinated hand-eye actions to set switches, adjust selector settings, and reposition control levers.

Reading and more broadly, the interpretation of graphical information are further important visual tasks. The approach plate and airport diagram are used as sources of information to support approach, landing, and taxi operations. Some of the information from these sources is purely textual; some involves the interpretation of annotated graphical information. Each of these visual operations is an important component of model behavior.

Vision plays a central role in the modeling of a pilot's multitask capabilities. The pilot's procedures for each of these visual activities has a priority associated with it. Within this framework, the vision system is a resource and the tasks that require its capabilities compete based on their assigned priority. As modeled, the scans of the windscreen and the individual flight deck instruments, the PFD, the HSI, and the SVS when it is present, are modeled as separate procedures, all with the same priority. The background scan pattern thus produced, moves smoothly to the windscreen and from instrument to instrument. The interval between scans for each are adjusted for each instrument as appropriate as the approach and landing progresses (e.g., the scan of the HSI becomes more frequent at flight path waypoints). At particular points in the approach, instruments may be dropped from the scan pattern (e.g., the scan of the HSI after the decision to land). Clark (1999) describes similar variations in purpose directed saccade patterns as first reported by Yarbus (1967). To date, it has not been necessary to adjust the priorities for the basic scan procedures.

The action to scan out the windscreen for the runway or to support coordinated hand-eye actions to set a selector or control lever operate slightly differently. These actions are examples of a decision to take an action *now* and hence, are invoked at a priority higher than the background scan procedures. The elected action takes place immediately and once the action has completed, the background scan procedures resume. Actions with a visual component that extend over time operate with another slight variation. Visually tracking the runway to support the sidestep maneuver operates similarly with respect to priority, but to accommodate the extended time duration it allows intervals at which background visual procedures can intervene.

Information obtained by the pilot models in their scan of the HSI and the SVS is representative of the information derived from the major flight deck instruments. The HSI is a rich information source used by the pilots in tracking their progress along the flight path in each of the scenarios. Some of the information is immediately symbolic: the aircraft's heading, the distance to the next waypoint, and the display scaling. Some is readily determined in symbolic form once a little graphical interpretation has been accomplished: the name of the next waypoint. Some is geometrically interpreted: that the changing heading is converging on the desired heading for the new flight plan leg. Lastly, the pilots recognize the current waypoint as the last one before landing and use this information in deciding when to terminate their scan of the HSI. Figure 4 provides a screen shot of the HSI display as NASA186 is traversing the leg to GOLET. Each of the basic information items can be seen in the figure. The form of the display is based on the HSI display described in the RNAV cognitive task analysis (Keller & Leiden, 2002a).

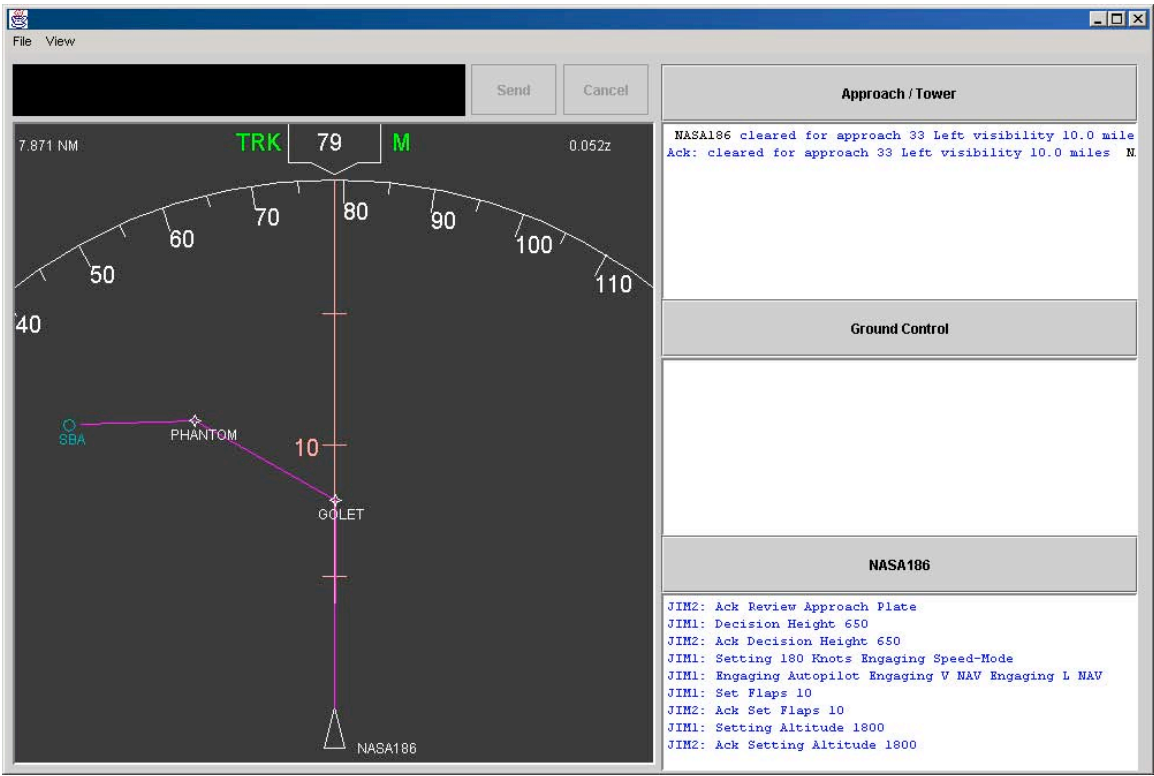


Figure 4. Horizontal Situation Indicator.

The scan of the SVS is particularly important to this research effort. What we now have in place, while sufficient to address the current scenarios, will need to be improved as we address scenarios that are more challenging and as we pursue the identification and mitigation of aircrew errors. For the present, the pilots readily derive heading, speed, altitude, and altitude rate as numeric quantities from the SVS much as they do from the PFD. As the airport comes into the display's field of view, the pilot's view of the runway is based on the distance to the runway. The SVS-equipped scenarios modeled to date are in IMC conditions with an 800-foot cloud ceiling. The pilots are able to track the runway using the SVS before the aircraft breaks out of the cloud cover.

2.5.4 A Distributed Model of Working Memory

On the flight deck, each aircrew member will typically have several tasks at various stages of completion. In pursuing the completion of each ongoing task, there will usually be several goals and procedures concurrently active. As the aircraft proceeds along the flight leg to PHANTOM in IMC with an 800-foot cloud ceiling, at least two of the captain's tasks are concerned with current altitude. For our example, let us say that the altimeter is currently reading 1021 feet afl. The more immediate task will be concerned with monitoring the descent to the target altitude of 1000 feet for the flight leg to PHANTOM. The second task concerned with tracking the current altitude, has as it goal, the decision to land to be made as the aircraft descends through the decision height of 650 feet. There is other work in process, but these two tasks are sufficient to suggest how the altimeter reading is processed in the model.

For all scenarios, an altitude reading is available from the PFD. For those scenarios that include an SVS, the SVS provides a second source for an altitude reading. The captain and the first officer each periodically scan the PDF (and separately, the captain scans the SVS when present). In our example, we will follow the captain's reading of the PFD. The scans each provide, at minimum, the heading, altitude, and speed in numeric form. In the vernacular of the model, upon completion of the reading of the PFD the value of "1021" for the altitude is "published"—that is, the labeled value is part of an attitude scan message that is the product of "scanning" the PFD (or the SVS).

In our example, at least two of the captain's procedures "subscribe" to attitude scan messages, the first concerned with monitoring the descent to the target altitude of 1000 feet for the current flight leg to PHANTOM, and the second preparing to make the decision to land. The active scan procedures for both the PFD and the SVS publish attitude messages, which in effect, "wake" the procedures that are subscribed to the message type. The output of a proactive visual process triggers cognitive processes requiring that output to facilitate their next actions.

From the captain's perspective, the value "1021" is processed in a unique manner in each procedure. For the task of monitoring the descent to the target altitude, "1021" immediately becomes "I need to attend carefully to the aircraft's altitude over the next few seconds to assure that we level off at 1000 feet." For the decision to land task, "1021" becomes "I've got a little time, but I should be breaking out of the cloud cover shortly and I then need to acquire to the runway to support the decision to land." In each case, the published altitude value of "1021" is immediately and significantly reinterpreted by the subscribing procedures. The value "1021" is just an intermediate value in a rapid succession of transformations.

The model's publish-subscribe protocol is designed to mirror the movement of information through the brain's visual centers and on to cognitive centers for further interpretation, action planning, and action execution. The model glosses over the many processing steps in the visual center that produce the symbol "1021," but then more faithfully represents the further processing of "1021" that led to the captain's actions related to assuring that the aircraft levels off at 1000 feet and to preparations to look for the break in the cloud cover and the sighting of runway 33L.

"1021" is certainly a working memory item, but it is simply one local value at one stage in a multi-branching process of transformations and interpretations starting in the visual system and moving through cognitive areas and then to motors areas that drive the resultant actions. Working memory items need a home in the architecture for a human performance model, but they cannot be properly captured in a single box in an architecture diagram. In the D-OMAR model, we posit that working memory, as we have just seen, is widely *distributed* across brain centers. Moreover, we assert that these working memory items do not have a separate existence as database items, but rather are each encapsulated by local processes, the procedures that operate on them at each stage of their migration, transformation, and interpretation.

3 Findings

Aircrew model development and the successful execution of the baseline and SVS-equipped part-task simulation scenarios led to findings related the use of the SVS as a second attitude instrument and findings related to the development of more capable human performance models.

3.1 Successful Scenario Execution

The D-OMAR aircrews, as expected, readily accomplished the five modeled scenarios. For the baseline scenarios in VMC and IMC conditions at SBA, the modeled aircrews successfully executed the approach and landing using RNAV procedures much as the human subjects did in the part-task simulation. The story was much the same for the nominal approach in IMC conditions using the SVS-equipped flight deck. When on the baseline VMC approach and in the SVS-equipped IMC approach, the tower controller requested that the aircrew side step from SBA runway 33L to the closely parallel runway 33R, the captain instructed the first officer to accept the request and then successfully executed the side step maneuver to runway 33R.

Actual performance of the aircrew models was reviewed at several levels of detail either during scenario execution or by reviewing data collected during a simulation run. A time-tagged on-line trace (see Figure 2) tracked the aircrew's conversation on the flight deck as well as the exchanges with the controllers managing the airspace. The trace also tracked flight deck actions taken by the aircrew that followed from this discourse. These traces confirmed that aircrew performance followed the procedures laid out in the Keller and Leiden (2002a) cognitive task analysis. A more detailed view of aircrew performance was reviewed using the Gantt-style display (see Figure 3) of goal and procedure execution for the captain and first officer. This display was used to review and evaluate aircrew performance at the task level by examining the timeline for goal, sub-goal, and procedure execution leading to task completions.

3.2 The SVS as a Second Attitude Display has Workload Implications

The addition of the SVS display to the flight deck augmented the out-the-window view while at the same time providing much of the same functionality as the PFD. In our model, the captain used the SVS to view runway 33L while still in the cloud cover, but reverted to the out-the-window view once the runway came in sight. Interestingly, there were individual differences in the behaviors of the three subjects in the part-task experiment during the flight phase from decision height to landing. While subjects four and five made the expected use of the out-the-window view, subject three relied more heavily on the SVS using the out-the-window view for only five percent of the flight phase.

When the SVS was added to flight deck, the captain, as modeled, included both the SVS and PFD in the scan for aircraft attitude, speed, altitude, and altitude rate information. The basic scan then included the out-the-window view, the PFD, the SVS, and the HSI. One impact of the scan of the two flight deck instruments with an overlap in functionality was that less time was devoted to the HSI display and the navigation function that it supports. Upon further review, the same effect was found in human subject data for the part-task simulation. A finding identified in the modeling process, was further supported by human subject data from the part-task simulation.

In viewing the SVS, there is certainly the sense that some elements of the terrain are more important than others, and that this is situation dependent, changing as the scenario progresses. During a nominal approach to SBA 33L, the target runway, once in range, may well be the most important SVS display feature. However, this would quickly change in the event of a go-around, particularly at an airport such as SBA where there are several hills along the 33L go-around route. Similarly, in the case of the sidestep maneuver to parallel runway 33R, the focus of attention must shift from 33L to the new runway assignment. The temptation is to highlight the salient features in the SVS display and adjust the highlighting to newly salient features as the approach progresses or as the situation changes. The request from the controller to shift the landing to 33R is simply an auditory radio exchange and the pilot assumes manual control of the aircraft to execute the sidestep maneuver. Unfortunately, it is not clear how to include the SVS in the information transfer related to decision to shift the landing to runway 33R. The automation is not a party to the exchange. It receives no information from which to capture the decision to land on 33R and has no basis from which to act to highlight 33R as the new target runway. The decision is conducted behind its back. Data link is a possible path by which to make the privileged information more broadly available on the flight deck. It might then be possible to think about highlighting salient SVS display features.

3.3 Additional Scenario Complexity Can Identify Model Shortcomings

Using a second aircraft in the late reassignment scenario was a small step in adding to the realism of the scenario. It was enough to trigger the need to refine model behaviors in dealing with the conflict that arises when there is air traffic controller communication at the same time that an aircrew has critical actions that they need to coordinate on. The nominal behavior is simply to have the aircrew defer their flight deck conversation when the air traffic controller interrupts a conversation in progress. This basic rule broke down first in the O'Hare scenario when the first officer had to get an okay from the captain before accepting the exit taxiway offered by the approach controller. The captain "spoke through" the on-going air traffic controller exchange to provide the okay to the first officer. In the late reassignment scenario, the lead aircraft was approaching decision height as the tower controller was providing the trailing aircraft with the clearance to land. The lead aircraft's first officer was monitoring the descent to decision height and similarly had to "speak through" the air traffic controller conversation with the trailing aircraft to notify the captain of its approach. Including more complexity in the scenarios teased out flaws in the models and lead to model improvements.

4 Implications

For the five SBA scenarios executed by the D-OMAR aircrew models, the aircrews readily accomplished the approach and landing tasks using the baseline and SVS-equipped flight decks. While aircrew performance readily meet the requirements of the scenarios, the SVS as a second attitude display was found to have implications with respect to aircrew workload. With respect to the models themselves, greater scenario complexity was one means by which to stress model capabilities, identify their shortcomings, and drive the development of more capable models. As model capabilities improve, it will be possible to more effectively address complex aircrew issues in aviation safety challenges.

4.1 Two Attitude Displays or One?

The modeled aircrews and the subjects in the part-task experiments tended to spend less time attending to the HSI display when the SVS was available even in the early phases of the approach where they were principally monitoring their progress along the flight path. The presence of the SVS appeared to reduce the time allocated to the HSI even when the HSI was the information source most relevant to the current flight phase. A simple explanation might be that the aircrews had sufficient time to accomplish their navigation task and were simply using the HSI as required. Nevertheless, it is possible that there is a underlying problem.

On the SVS-equipped flight deck, the aircrews effectively have two attitude instruments. In scanning these two separate instruments, it appears that they may be drawn to spend more time attending to attitude-related information than would be necessary when using a single attitude-instrument configuration. With the dual instrument configuration, the habituated pattern may become an unduly complex two-instrument scan with attitude information derived from the two similar but not identical displays. In situations where time pressure is high, having two instruments from which to obtain required information can impose additional burdens on the pilot. Does the pilot stick with the time consuming habituated scan or try to save time by switching to a single instrument, but less practiced scan? Even electing to consider the options has a cost. Moreover, the extra cognitive effort required to switch to a single instrument scan may negate its potential inherent advantage. Changing a habituated two-instrument scan pattern when change is most difficult is an imposition on the aircrew that should be avoided if possible.

One potential solution is to consider a single attitude display combining PFD and SVS functionality. An SVS that has nominal PFD behavior as a fail-safe mode might, if feasible, be considered and explored. The habituated scan would then be a single instrument scan that readily avoids the choice of scan pattern that the two-instrument configuration can potentially impose, just when choice is most difficult.

4.2 Complex Scenarios Can Drive Necessary Model Improvements

The addition of a second aircraft in the late reassignment scenario proved to be enough to identify a subtle improvement needed in an aircrew model's handling of an ATC communication with another aircraft at a time of high workload—the first officer "spoke through" the ATC communication to let his/her captain know they were approaching decision height. Scenarios that are more complex stress our models, but more importantly drive the development of models that provide better insight into pilot behaviors. Increased workload can negatively impact aircrew performance. Increased workload for human performance models may well promote the development of better models—models that help us to understand resulting changes in aircrew performance, enable us to identify the errors that might be the product of high workload situations, and provide a setting in which to evaluate error mitigation strategies.

5 Lessons Learned

D-OMAR provided excellent support for developing the five initial SBA approach and landing scenarios. The aircrew models that executed the ILS landings at O'Hare for last year's study proved to be readily extendable. The RNAV approach, as detailed in the cognitive task analysis (Keller & Leiden, 2002a), required the implementation of broad range of new goals and procedures, but it was relatively easy to accomplish that within the framework for the human performance models established for the O'Hare scenarios. As the RNAV procedures were developed, they also shared many of the goals and procedures for the ILS approach. As the changes were made, they did at times impact the ILS code and it was necessary to take additional steps to insure the continued operation of the O'Hare scenarios.

In effect, we now have a working RNAV approach and the ILS procedures are more complete.

Constructing the SBA airport model was done using data structures developed for the O'Hare model. Information on Santa Barbara Municipal Airport was obtained primarily from the AIRNAV.COM web page for the airport. The web page included latitude-longitude information for the endpoints of the runways, information on radio frequencies for the controllers, and a link to an airport diagram. Information on runway signage was not available, and was constructed to enable taxi operations similar to those used in the O'Hare scenarios. The AIRNAV.COM pages also provided information on the radio navigation aids for the approach to SBA 33L. The navigation aid information obtained included location latitude and longitude, and radio frequencies. The availability of an airport physical description database would certainly help the process of constructing an airport model. The developers and maintainers of AIRNAV.COM have provided an important first step in this direction.

The NASA part-task simulation data has proven to be a very valuable resource with much still to be learned. In particular, the fixation sequence data files and the eye tracker video tape provide a level of detail in instrument scanning that our models ought to more accurately represent. For the present, the models look at an instrument and read the instrument's data items in a single pass. The eye tracker fixation data videotapes suggest that the pilots selectively and repeatedly scan individual items within a display before moving on to the next display. Clark (1998) reports similar patterns of saccade sites being visited and revisited in the short term. A number of recent studies (Wickens, Xu, Helleberg, Carbonari, & Marsh, 2000; Diez, Boehm-Davis, Holt, Pinney, Hansberger, Schoppek, 2001; Hüttig, Anders, & Tautz, 1999; Anders, 2001) have used eye tracking to examine pilot scan patterns. To the extent that this data is made available, there will be more and more data available to tell us *what* the pilots' scan patterns are. As the developers of pilot models and the theory that underlies our models, we must seek to explain *how* basic human capabilities operate together to yield these scan patterns. Our models will better represent pilot performance to the extent that these behaviors are better explained.

Our long-term goal remains to make use of the understanding of pilot behaviors as represented in human performance models to explore the means to "reduce accidents by mitigating system-wide accident precursors." Aircrews, the pilots in the part-task experiments and the D-OMAR pilot models in the SBA scenarios, readily make appropriate use of the SVS. The scenarios as modeled, add complexity to the scenarios as executed in the part-task scenarios. We would like to build further complexity into the scenarios, refine aircrew procedures for the use of the SVS (possibly eliminating the PFD for some scenarios), and run a series of trials with the aim of probing modeled pilot behaviors for potential benefits as well as errors that might occur on the SVS-equipped flight deck. For human performance shortfalls that lead to errors, we would like to examine approaches to mitigate those errors.

6 References

Anders, G. (2001). Pilot's attention allocation during approach and landing: Eye- and head-tracking research in an A330 full flight simulator. In *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH.

Anderson, J. R., & Lebiere, C. J. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bargh, J. A, & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462-479.

Clark, A. (1999). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Corker, K. M., & Smith, B. R. (1993). An architecture and model for cognitive engineering simulation analysis: Application to advanced aviation automation. In *Proceedings of the AIAA Computing in Aerospace 9 Conference*. San Diego, CA.

Deutsch, S. E., & Pew, R. W. (2001). *Modeling human error in D-OMAR*. (BBN Report 8328.) Cambridge, MA: BBN Technologies.

Deutsch, S. E., & Pew, R. W. (2002). Modeling human error in a real-world teamwork environment. In *Proceedings of the Twentieth-fourth Annual Meeting of the Cognitive Science Society (pp. 274-279)*. Fairfax, VA.

Deutsch, S. E. (1998). Interdisciplinary foundations for multiple-task human performance modeling in OMAR. In *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society*. Madison, WI.

Deutsch, S. E., & Adams, M. J. (1995). The operator model architecture and its psychological framework. In *Proceedings of the 6th IFAC Symposium on Man-Machine Systems*. Cambridge, MA.

Diez, M., Boehm-Davis, D. A., Holt, R. W., Pinney, M. E., Hansberger, J. T., & Schoppek, W. (2001). Tracking pilot interactions with flight management systems through eye movements. In *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH.

Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. New York: Basic Books.

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness*. New York: Basic Books.

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.

Huettig, G. Anders, G., & Tautz, A. (1999). Mode awareness in a modern glass cockpit: Attention allocation to mode information. In *Proceedings of the 10th International Symposium on Aviation Psychology*. Columbus, OH.

Keller, J., & Leiden, K. (2002a). *Information to support the human performance modeling of a B757 flight crew during approach and landing: RNAV.*

Keller, J., & Leiden, K. (2002b). *Information to support the human performance modeling of a B757 flight crew during approach and landing: SVS addendum.*

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence, 33,* 1-64.

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30,* 583-598.

Meyer, D. E., & Kieras. D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104,* 3-65.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action*. London: Lawrence Erlbaum.

Posner, M. I., Peterson, S. E., Fox, P. T., & Raichle, M. E. (1988). Localization of cognitive operations in the human brain. *Science, 240,* 1627-1631.

Wickens, C. D., Xu, X., Helleberg, J. R., Carbonari, R., & Marsh, R. (2000). *The allocation of visual attention for aircraft monitoring and avoidance: Baseline measures and implications for freeflight*. Savoy, IL: University of Illinois.

Williams, D., Waller, M., Koelling, J., Burdette, D., Doyle, T., Capron, W., Barry, J., & Gifford, R. (2001). *Concept of operations for commercial and business aircraft synthetic vision systems, Version 1.0.*

Yarbus, A. F. (1967). *Eye movements and vision*. New York: Plenum Press.

Using an Integrated Task Network Model with a Cognitive Architecture to Assess the Impact of Technological Aids on Pilot Performance

Christian Lebiere, Rick Archer, and Dan Schunk
Micro Analysis and Design
Boulder, Colorado

Eric Biefeld
Carnegie Mellon University
Pittsburgh, Pennsylvania

Troy Kelly and Laurel Allender
U.S. Army Research Laboratory
Adelphi, Maryland

Introduction

NASA's System-Wide Accident Prevention Program (SWAP) has funded a number of different cognitive modeling approaches with the goal of developing techniques for predicting opportunities for human error in system designs and operational procedures. The project has been named the NASA Human Performance Modeling (HPM) project, according to Foyle, Goodman, and Hooey [2003]. Each of the modeling teams is developing a different approach to the same modeling problem. The current focus is on commercial aircraft approach and landing scenarios and the tasks the pilots must perform. The modeling compares pilot procedures using current technologies with procedures using augmented displays such as a synthetic vision system (SVS).

Data to characterize the performance of pilots during four different approach and landing scenarios was collected by NASA via pilot participation in a part-task simulator [Goodman, Hooey, Foyle, and Wilson, 2003]. The four scenarios consisted of 1) a vectored approach, 2) a late reassignment of runways for the landing, 3) a missed approach, and 4) a mismatch of the terrain shown in the SVS and the out the window view from the cockpit.

One of the modeling teams consisted of personnel from Micro Analysis & Design (MA&D) and Carnegie Mellon University (CMU). This paper describes the modeling approach used by the Micro Analysis and Design/Carnegie Mellon team and its application to one of the four scenarios.

Description of the Modeling Effort

The approach that was used by the MA&D and CMU team to perform the approach and landing modeling was an integration of the Improved Performance Research Integration Tool (IMPRINT) and the Adaptive Control of Thought – Rational (ACT-R) cognitive architecture. It was agreed by project personnel from NASA and each of the modeling teams that the first scenario for all of the modeling teams would be a late reassignment of runways. There have been several sources of data for the modeling effort. The primary data source was a Cognitive Task Analysis (CTA) that was conducted as a separate effort in support of all of the modeling teams. The CTA was supplemented by a number of published papers and other background documents. In addition, videotapes of pilots flying approach and landing tasks in a NASA part-task simulator were provided. Following is a brief description of the two integrated modeling tools and what each provides for the integration.
The results of the current year of the NASA HPM project from MA&D is a simulation model of an aircraft making its final approach and landing into an airport. The specific scenario that was chosen for this first effort was of an approach with a late reassignment of runways for the landing. The simulation model was built using the IMPRINT simulation tool [Archer and Allender, 2003]. The model built specifically represents an aircraft and its environment. Currently this environment includes the altitude at which the ground and runway can be seen from out the window and the air traffic control (ATC) communications. The model has been designed for more environment variables to be added as required. For the simulation of the aircraft, IMPRINT represents the autopilot as well as the physics of the aircraft. These aspects include the aircraft's location in time and space, its deceleration, descent rate, and all physical changes in the aircraft including its landing gear, flap settings and air brakes. The model also includes the controls and displays of the aircraft including all autopilot functions. Represented in the model are the mode control panels, the primary flight display, the navigational display, and an out-the-window view. The model also handles all communication between the aircraft and air traffic control. With these controls and displays, the model is able to simulate how a plane will react in its environment when these controls and displays are manipulated. Currently the simulation is set up to work with a VNAV Path autopilot setting as

required for this first effort, but the model is capable of utilizing the other types of autopilot (e.g. Glideslope and Localizer) for future analysis. In order to perform successful analyses, this model requires an outside data source to act as the pilot (i.e. human in the loop or cognitive modeling software). This data source will then issue look and manipulate commands to the controls and displays of the model as required to perform the approach and landing tasks. The simulation model will terminate when the pilot switches off the autopilot for the manual portion of the landing.

In this simulation, a model of the pilot was developed using the ACT-R cognitive architecture. Following the practice of decomposing complex behavior into a set of unit tasks, the ACT-R model is composed of a set of goals, together with the procedural and declarative knowledge necessary to solve those goals. The top-level goal is essentially a monitoring loop that repeatedly sets subgoals to check the settings of the various controls. Each of these subgoals typically requires acquiring the value of one or more environmental variables (e.g. speed, altitude, etc) by reading the instruments or looking out the window. A decision is then made as to what the desired control value is given those readings. If that value is different from the current control value, the appropriate action is performed to change that value. Decisions are made using either declarative or procedural means. For procedural control, a production rule is applied that supplies the control value given the environmental readings. This type of decision best captures crisp, symbolic decisions relying on precise values provided by instruments (e.g. "set flaps to 15 when speed is 200 knots"). For declarative control, instances are defined in declarative memory linking environmental readings to control values. Given a particular condition, the most relevant instance is retrieved from memory using a similarity-based partial matching mechanism, and the control value extracted from it. Multiple memory instances can also be retrieved using a mechanism called blending and a consensus control value extracted that best satisfies the set of instances. This control is similar to that provided by neural networks and best describes approximate, iterative adjustments as practiced in out-the-window flying.

The Imprint Model

The backbone of the IMPRINT simulation is a series of subnetworks representing the different aspects of the simulation (see Figure 1). When the IMPRINT simulation is started, the model starts in subnetwork 2, the Start/End Landing network. In this network all the initial communication takes place between ACT-R and IMPRINT. From the Start/End Landing subnetwork, the simulation starts the ACT-R Thought subnetwork, the Aircraft Environment subnetwork and the communication subnetwork.

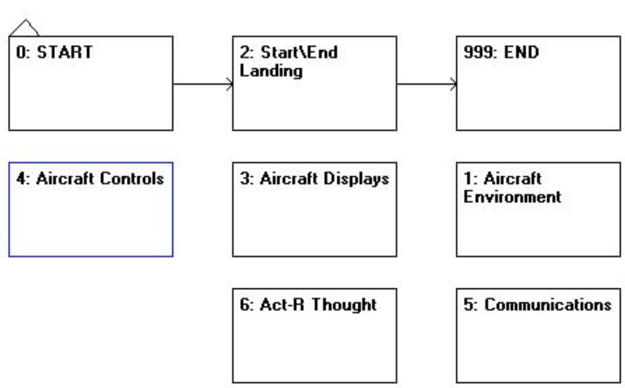


Figure 1: Main IMPRINT Network

When the communication subnetwork is started, IMPRINT will schedule random communication tasks between the pilot and other pilots as well continually check for scheduled communication between the pilot and air traffic control. Inside the ACT-R Thought subnetwork, the main communication between ACT-R and IMPRINT takes place. With this communication, ACT-R decides what controls and displays to look at or manipulate. The Aircraft Controls and Aircraft Displays subnetworks represent the various controls and displays of the aircraft. In the Aircraft Environment subnetwork, IMPRINT updates the current state of the aircraft, its controls, its displays, and the environment itself every .1 seconds. This is a continual action until the end of the simulation.

Figure 2: Aircraft Displays Subnetwork

An aircraft's displays are further broken down in the Aircraft Displays subnetwork (see Figure 2). In this subnetwork, the displays are separated into the main displays of the aircraft, specifically, the primary flight display, the navigational display, and the out the window display. It is important to point out that while there is no physical display specifically for out the window, the pilot has the ability to look out the window in order to perform such actions as seeing whether the ground is in sight or what the flap settings currently are at. The final level of the Aircraft Displays Subnetwork is the actual displays represented by tasks in the IMPRINT simulation. Figure 3 shows how each part of the Primary flight display is represented by a single task. During the IMRINT simulation, ACT-R will communicate to IMPRINT which display it would like the value for. IMPRINT will then start the task for the appropriate display and return that display's current value. The information is then communicated back to the ACT-R Thought subnetwork located in the main IMPRINT network (see Figure 1).

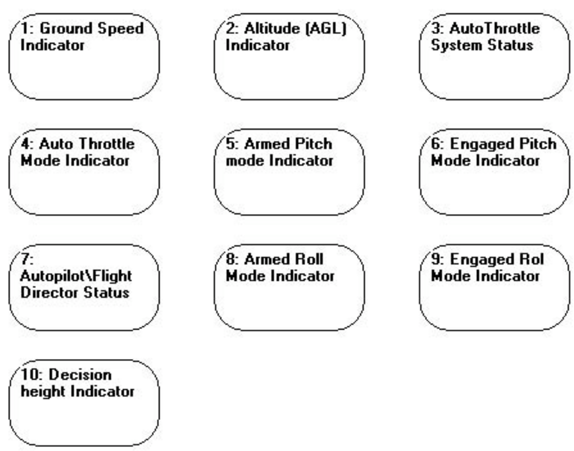


Figure 3: Primary Flight Display Subnetwork

The ACT-R Model

The ACT-R model cycles through six decisions. These decisions are the setting of the flaps, altitude, speed, lowering of the landing gear, setting of speed brakes, and engaging and disengaging the autopilot. The model keeps cycling through these decisions until it either turns off the autopilot or misses the runway.

The model uses two buffers, which mediate between ACT-R and the world (IMPRINT). The first buffer is LOOK. If ACT-R wants the value of some instrument or the setting of a dial, it uses the LOOK buffer to pose a query. The ACTION buffer is used to change the setting of a dial. The contents of both of these buffers are communicated to IMPRINT which then updates the LOOK buffer with the value stored within IMPRINT.

Most of the decisions are based on either the aircraft's speed or position. The Set-Flaps decision is based on the speed of the aircraft. The first decision is to look at the speed indicator. ACT-R then matches this speed against its speed flaps value in memory. After the flaps setting is retrieved, ACT-R looks at the current setting on the flap dial. If it is the same then it just goes to the next decision, in this case, dial altitude. If it is different it uses the ACTION buffer to update the flap setting.
The altitude settings depend on which waypoint is next in the approach phase. Once again, ACT-R uses the LOOK buffer to "read" the waypoint indicator. It then "remembers" the correct altitude for the waypoint and if it is not already set, it communicates to IMPRINT to dial in the new altitude. If it is working correctly then the altitude should only be changed once for a waypoint position.

The speed setting depends on the distance to the airfield. However, there is not a single reading which directly gives this data. Instead, ACT-R looks at the waypoint and then looks at the distance to waypoint. ACT-R then recalls the range of this waypoint to the airfield and computes the distance to the airfield. After this more complex set of productions, ACT-R then retrieves and, if needed, sets the speed in the same style as the previous decisions.

The other decisions are done in a slightly different style. These are more like rule-based productions. Instead of being based on a set of memorized relations, the productions directly encode the decision point. For example, the abort-landing production checks to see if the distance to the runway is zero. If it is, it signals an abort and pops all the goals.

The decide gear production checks if the distance to the runway is less than or equal to 15.0 miles. If it is, it lowers the landing gear. There is currently no production for raising the gear.

The decision for the speed brakes is to compare the current speed to 145.0 knots. If the speed is less, the brakes are turned off. If the speed is greater than 145, the brakes are turned on. Like the other decisions, before performing the action a production checks the current setting and performs the action if the decision is not equal to the current setting. Unlike the landing gear, the current model could turn off the speed brakes, but given the current scenarios this never happens.

The decision to disengage the autopilot depends on the visibility of the runway. ACT-R uses the out-the-window values to determine if the runway is visible. If it is then the next production will disengage the autopilot.

When control returns to the top-level decision cycle, if the autopilot has been disengaged, ACT-R ends the simulation and signals success. If the autopilot is engaged, ACT-R cycles back to the set flap decision.

A separate production responds to communications. The only effect it has is to take time away from the top-level control loop. This can cause the model to "miss" the runway.

Act-R/Imprint Protocol for the NASA HPM Project

In order to successfully have communications between IMPRINT and ACT-R a protocol was set up so that each program responded correctly to the other program's request. In this project IMPRINT acts as the aircraft and the environment affecting the aircraft. ACT-R acts as the pilot of the aircraft and decides which displays to look at and which controls to manipulate. The first step in creating this protocol was to develop a simple generic protocol that all ACT-R/IMPRINT links will use as a basis for developing intercommunications with each other. This generic protocol has three main areas: starting a model, ACT-R actions to be performed that take time, and ending a model.

When an IMPRINT model that requires communication with an ACT-R model begins executing, an initial communication must take place between ACT-R and IMPRINT. This is done to ensure that both programs are synchronized in all areas of the model. This allows ACT-R to reset its current data as well as check to make sure that both IMPRINT and ACT-R are running the same version of the model. In the protocol, when the initial communication is made, the version numbers of both IMPRINT and ACT-R are exchanged; if there is a mismatch then the model will end.

During the execution of an IMPRINT model, decisions and thought processes will need to be made by ACT-R that will take time. This time needs to be included inside the IMPRINT model, which holds the master simulation clock. The biggest obstacle with this communication is that during the time it takes for ACT-R to perform a thought or make a decision, an event in the IMPRINT simulation may occur. This event may cause ACT-R to perform differently (e.g. a person driving a car is thinking about changing the radio station when his/her tire blows out causing them the change their thought process from what radio station to listen to, to controlling the automobile). Likewise, ACT-R may decide to do nothing at the time but calculates a decision that may be made in the future (e.g. a driver of a car is deciding whether to change lanes or not, initially the driver decides to do nothing but ten seconds later decides to switch lanes). One thing to note is that the event may NOT cause ACT-R to perform differently (e.g. a person driving a car makes the decision to decelerate the car when the sun gets in their eyes, their decision stays the same). In the protocol, IMPRINT sends to ACT-R the current time of the model as well as the time in the model that the next IMPRINT event will occur. ACT-R, after performing its calculations, will reply back to IMPRINT with the new current time, the time that it will have another thought, and the action it has decided to perform. From this interaction, IMPRINT will decide what the next event will be as well as how much time it took ACT-R to make its decision. This interaction will occur many times in the model as the simulation is executed.

The final generic protocol is a communication between IMPRINT and ACT-R that allows both programs to end correctly. The final event in the IMPRINT model is a communication between IMPRINT and ACT-R where IMPRINT sends over information to let ACT-R know that the IMPRINT simulation is ending and that this will be the final communication between the two programs. In this protocol, IMPRINT will send over a flag to ACT-R. ACT-R will then acknowledge that the flag has been received and end its model. IMPRINT will then end its simulation.

Though the generic protocol was used as the basis, there were some changes to the basic interactions between IMPRINT and ACT-R as well as some additional interactions. The only generic protocol changed was the interaction between IMPRINT and ACT-R that occurred during the run. The added interactions include return information to ACT-R from IMPRINT, pilot/air traffic control communication, and some initialization of the model.

The main change to the generic protocol was to get more specific as to what actions the pilot was performing onto the aircraft. In this interaction, ACT-R is not only responsible for informing IMPRINT how much time it took for the pilot to make a decision on what to do, ACT-R also had to inform IMPRINT on what actions to take. The actions could be one of four things; the pilot could manipulate a control, look at a display, communicate with ATC or decide to do nothing. In informing IMPRINT of the action, ACT-R needed to provide IMPRINT with this specific information: what action to perform, what display or control the action was performed on, and the new setting of the control (if applicable). If the pilot was looking at a display, the action would default to a look action, but because the controls can be switches, buttons, or knobs, the correct action also needs to be performed on the control and there could be a possible setting (e.g. altitude setting). If the action was to communicate with ATC, ACT-R would also need to let IMPRINT know the type of reply the pilot was giving to ATC.

Once ACT-R makes a decision as to what type of action to perform and where to perform it, IMPRINT needs to perform that action on the aircraft and update ACT-R as to the current state of the display/control. A real world example of looking at a display would be if a pilot decides to look at the current airspeed. In the real world, a pilot would look at the display and the pilot's visual buffer would be instantaneously updated to the current speed. In the IMPRINT/ACT-R simulation, and when ACT-R (acting as the pilot) wants to look at a display, IMPRINT (acting as the aircraft and its displays) will need to return the value of that display to ACT-R to update ACT-R's memory. In the model, IMPRINT would change the simulation clock time to include the time it takes for ACT-R to make the decision to look at a display and return the display's current value. While this interaction will take time to perform in real world time, the simulation clock will not advance (and this interaction will be instantaneous). If a pilot wanted to manipulate a control (such as an altitude dial) then the pilot would take time to decide upon the new altitude, dial in the new altitude, and his visual buffer would be updated to the new altitude number. In the IMPRINT/ACT-R model, ACT-R would make the decision to change the altitude dial to a new value. IMPRINT would then advance the clock the time it took to make the decision to turn the dial and advance the clock the time it took to advance the dial. IMPRINT would then communicate to ACT-R the dial's new value. Once again this will take up real world time, but in the simulation the update would be instantaneous.

When a pilot is flying an aircraft, communications between the pilot and ATC will occur at various intervals. Communications fall into three different categories:

a. Communication to which the pilot should respond
b. Communication to which the pilot should listen
c. Communication that the pilot should ignore

During the simulation execution, various communications take place between ATC and the pilot. When a communication event occurs, IMPRINT will interact with ACT-R informing ACT-R that an ATC communication has occurred as well as the type of the communication. ACT-R then makes the decision during the course of the simulation as to if and when it should reply back to ATC in the form of an action.

Appendix A contains details of the IMPRINT communication protocol with ACT-R. Appendix B contains details of the ACT-R communication protocol with IMPRINT.

Findings and Implications

The model has many potential parameters, but we can aggregate them into five main ones, represented in figures 4-8 below. Each of these parameters corresponds to a separate module in the ACT-R architecture. Variations in those parameters can correspond to individual differences as well as changes in equipment or procedures. Four of the parameters are latencies (in seconds) which represent the time for the pilot(s) to perform procedural, visual, motor and auditory actions. Anderson and Lebiere, [1998] state that the other is ACT-R's Activation Noise value, which is a measure of the stochasticity of the model's decision-making (see "The Atomic Components of Thought" for details). The first parameter is "visual speed", which represents the mean time for the pilot to look at an instrument and perceives its value. The next is "manual speed", which is the mean time it takes to perform an action such as dialing in a new setting. The next is the "procedural speed" parameter which represents the time for a production rule to fire. The last is "aural speed", which represents the time that the pilot spends listening to a communication before replying. The visual, manual and aural speed parameters are scaling factors, which vary depending on the precise perceptual, motor, or communication act. These values specify a random distribution to represent adequate between- and within-subjects variability.

To test the model's sensitivity to these parameters, their values were varied over a range of possible quantities. Since all five parameters had similar values we used the same test range for each parameter.

Figures 4-8 below show the influence of the different pilot psychometric parameters on performance and their implications for cockpit design and procedures. The performance measure is performance correct landings as estimated by Monte Carlo runs of 20 samples for each parameter combination. The percentage of correct landings for default parameters was about 90%.

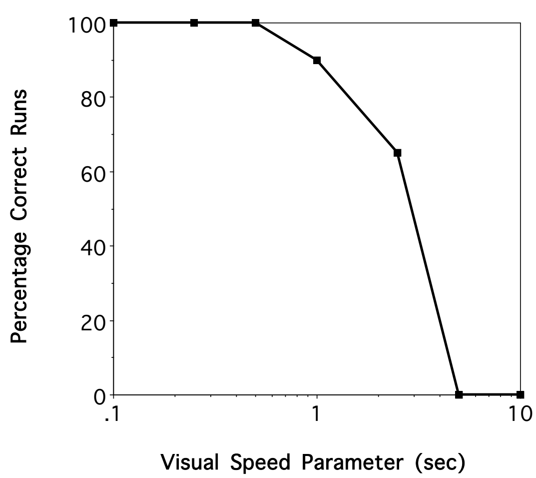


Figure 4. Sensitivity to visual speed

Figure 4 shows that pilot performance is very sensitive to the speed of visual shifts. The default value is 1 second, which corresponds to 90% correct landings. Improving the visual speed parameter by a factor of 2 yields perfect performance, which emphasizes the benefits of visual aids such as SVS that can improve perceptual performance by combining multiple instruments onto a single display, thereby reducing the time taken by attention shifts between widely separated instruments or areas such as out-the-window scene. Conversely, impairments in visual speed lead to a rapid deterioration in performance.

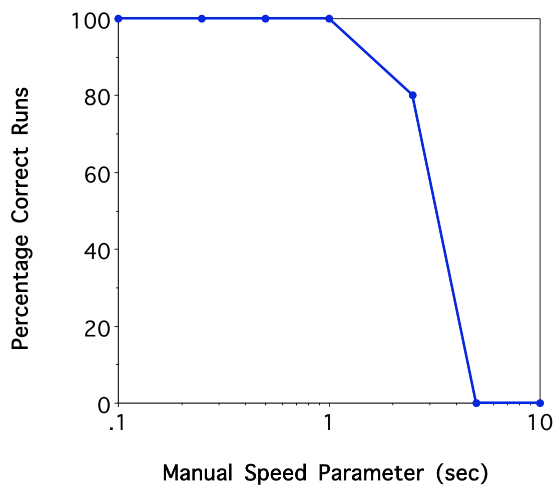


Figure 5. Sensitivity to manual speed

Figure 5 shows that pilot performance is also very sensitive to the speed of manual operations. The default manual speed is 2.5 seconds per action. Performance drops catastrophically when actions take significantly longer to perform, leading to total failure for a doubling of the action time. This factor supports the concept of division of labor between the pilot flying the aircraft and the pilot not flying, with one in charge of monitoring and decision-making and the other in charge of actually performing the actions.

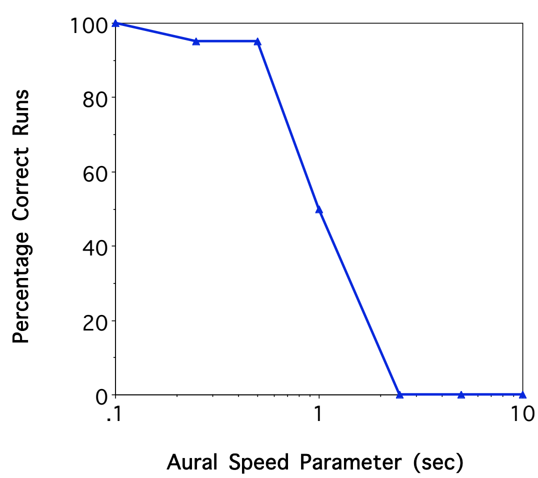


Figure 6. Sensitivity to communications

Figure 6 shows that performance is highly sensitive to the overhead of communications. The default time to listen to a communication and decide whether it is relevant to performance is 0.5 second, which is quite small. However, even when reduced to still smaller values, a small risk of failure remains. The impact of communications is hard to eliminate because of the random nature of their temporal distributions. Communications can occur at any time, including at critical moments in the control loop when even a small temporary disruption in attention can lead to failure. Increases in the number of communications or in the average duration of each communication (one can view the aural speed parameter as reflecting either) leads to a very rapid deterioration in performance.

Figure 7 shows that performance degrades gradually with the activation noise that controls the stochasticity in the memory retrieval of decision instances. Interestingly, further reducing activation noise from its default value of 0.1 does not lead in the elimination of errors, because activation noise has been shown to serve a useful purpose in preventing systematic errors. The most promising remediation for this degradation is training to increase the number of decision instances (practice) and/or making the decision task more procedural.

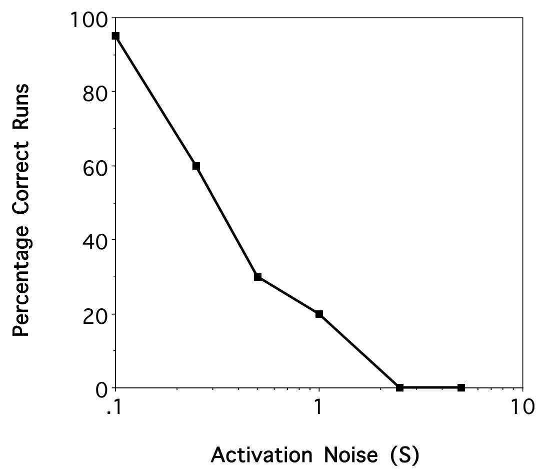


Figure 7. Sensitivity to decision consistency

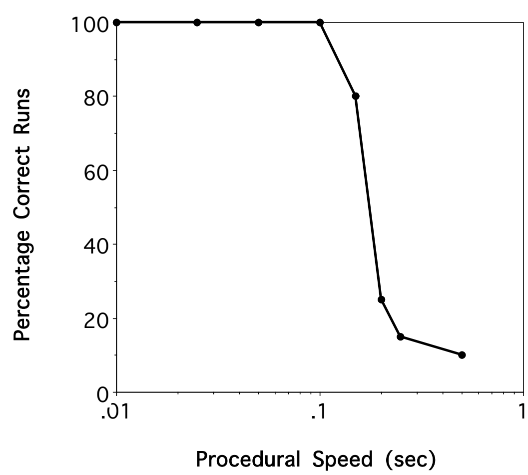


Figure 8. Sensitivity to procedural speed

Figure 8 shows that, because of the central nature of productions, performance degrades very sharply with the speed of the production cycle. Fortunately, there is a safety factor of 2 from the default production cycle latency of 50 msec before performance degradation occurs. This primarily reflects the fact that the production cycle is one to two orders of magnitude faster than the other ACT-R modules, and that performance is therefore more limited by the speed of the other modules than by the production cycle itself.

At this point, running the model with and without synthetic vision technology will produce very similar results since both conditions produce clear daytime vision of the terrain and runway. The important difference is that the average perceptual time is significantly reduced, on the order of a Look parameter value of about 0.5. This reduction occurred primarily because shifts of attention are greatly reduced because of the integrated display. This puts SVS operation in the range of safe, successful performance rather than the break point pictured above for conventional systems. We still need to more finely model additional savings achieved from the SVS display.

Future Directions

Both the IMPRINT and the ACT-R models are currently being expanded to include all of the controls, displays, and autopilots in the B757 aircraft. This will allow us to model a number of different baseline and SVS flight scenarios. In addition, we have included the capability to quickly model other new cockpit technologies and their impact on pilot error and aircraft safety. In the upcoming year, we plan to expand our current models to include multiple cockpit scenarios with and without augmented displays.

Verification and Validation

Given the dynamic nature of the timing of the decisions and actions captured in this modeling effort - both in terms of the internal cognitive environment as well as the external airport environment - the validation of the modeling work focused on the times associated with approaching and landing the aircraft. The validation process consisted of checking the predicted times associated with the aircraft activities against the times measured for the videotaped scenarios. The model predictions were based on inputs gathered from subject matter expert (SME) reports. The SME reports were used to estimate timing for all aircraft phases. These data were then fed into the IMPRINT model as environmental variables, which then were passed to the ACT-R model as cognitive variables. The interaction of these two models produced final model times as well as patterns of errors. The times were checked against the times measured from the supplied videotapes. Although no statistics were computed, the model outputs compared favorably with the times from the videotaped scenarios.

Another source of validation is low-level behavioral data like duration and aggregate percentages of dwell times on various parts of the display as provided by eye-tracking equipment. Following a top-down modeling methodology, we initially chose to focus on the functional aspect of the model and the sensitivity of its overall performance to various architectural parameters, as described in a previous section, and to higher-level behavioral measurements as mentioned in the previous paragraph. While the model makes precise predictions about the amount and duration of times spent acquiring information from the environment (e.g. reading the altitude), it does not contain specific assumptions regarding which source that information is acquired from, e.g. from a traditional instrument or from the symbology of the SVS display. Such assumptions are required to make specific predictions regarding eye movements. However, as indicated by analyses of the eye tracking data, fundamental individual differences exist even within small subject populations (e.g. 3 pilots). Thus, one pilot might choose to rely primarily on traditional instruments, using the SVS system only to confirm their readings, while another might choose to rely on the SVS as its primary integrated source of information. To model both of these (as well as other) strategies, the model would have to be flexible enough to incorporate a learning component that would allow the model to settle on either of those strategies as a valid stable point in its search for information. We intend to refine our model to incorporate learning of information-gathering strategies in order to generate eye movement predictions. This model will rely upon the Bayesian mechanism built into ACT-R to learn the utility of production rules, such as those selecting where to look to gather a piece of information.

Lesson Learned

Our main advance in performance modeling consists of linking a discrete-event simulation tool such as IMPRINT to a cognitive architecture such as ACT-R. This combination works quite well in alleviating the shortcomings of each platform: the cognitive architecture provides a higher-fidelity representation of human performance than task network models, and the discrete-event simulation provides a transparent, scalable representation of the world to interact with the human performance models.

We learned that, in order to accurately model the behavior of the pilot and to be able to predict errors that the pilot will make, we need as much specific information as possible, especially in terms of the response of the aircraft to various commands. Ideally, we would like to hook our ACT-R model directly to the NASA part task simulator. In this way we could ensure that we present the model with the full rigor of the task and replicate exactly what the pilots did in the experiment.

References

Archer, S. and Allender, L. "New Capabilities in the Army's Human Performance Modeling Tool," *Proceedings of the Military, Government, and Aerospace Simulation (MGA 2001) Conference,* pg 22-27, editor Michael Chinni, 2001.

Anderson, J. and C. Lebiere (1998). The Atomic Components of Thought. Mahwah, NJ, Lawrence Earlbaum Ass.

Box, D. (1998). Essential COM. Reading, MA, Addison Wesley Longman, Inc.

Foyle, D.C., Goodman, A., and Hooey, B.L. (2003). An overview of the NASA Aviation Safety Program (AvSP) System-Wide Accident Prevention (SWAP) Human Performance Modeling (HPM) element. 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.

Fujimoto, R. M. (2000). Parallel and Distributed Simulation Systems. New York, NY, Wiley-Interscience.

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.

APPENDIX A
IMPRINT COMMUNICATION PROTOCOLS

Communications between ACT-R and IMPRINT are made using External Model Calls (EMCs). EMCs utilize Microsoft's Component Object Model (COM) capability to allow two or more programs to share information. When an IMPRINT model is loaded into Goalsaint, its discrete event simulator, a communication link is created between the Imprint model and ACT-R. Every time a variable is sent between ACT-R and IMPRINT, the software receiving the variable must send out an acknowledgement that the variable was received (e.g. When ACT-R receives a variable from IMPRINT, ACT-R must return to IMPRINT a value of one to acknowledge that ACT-R did in fact receive the variable from IMPRINT) During model execution, Goalsaint will execute model commands that send variable information to ACT-R. ACT-R will then take this information and make the appropriate decision as to what the pilot will decide to do next. The following is a list and explanation of each type of External Model Call made in the NASA HPM Imprint model.

EMC Definitions:

EMCStart

Purpose: Initialize contact between ACT-R and IMPRINT.

Vars Sent to ACT-R

VersionNum – The current version of the IMPRINT analysis
EMCFlag – Flag representing which EMC call is currently being made

Vars Returned to IMPRINT
ACT-RVersNum – The current version of the ACT-R analysis

EMCEnd

Purpose: End contact between ACT-R and IMPRINT

Vars Sent to ACT-R

ModelDone – Variable sent to ACT-R indicating the Model is ending
EMCFlag – Flag representing which EMC Call is currently being made

Vars Returned to IMPRINT

None – No variables are returned, this EMC call is just a message to ACT-R that IMPRINT will be ending the simulation and severing the communications link.

EMCAction

Purpose: Communication between ACT-R and IMPRINT, in which ACT-R as the pilot decides what control or display it will look at or manipulate.

Vars Sent to ACT-R

CurrentTime – The current simulation time
ResumeBy – The time at which an event in IMPRINT may occur
EMCFlag – Flag representing which EMC call is currently being made

Vars Returned to IMPRINT

CurrentTime – The current simulation time (Note: This time will include the time it took for ACT-R to make its decision, simulating the time it would take a pilot to make a decision)
ResumeBy – The time at which ACT-R may make another decision
ActionSet – The setting value of the control being manipulated, not all controls have specific settings (e.g. An altitude dial needs to have an associated dial value, and autopilot button just needs to be pushed)
ActionCon – What control/display is being looked at/manipulated
Action – What action is being performed on the control/display (e.g. look, push button, switch switch, etc)

EMCComm

Purpose: At various times during a flying simulation, there will be various communications being broadcasted, either randomly or at certain periods of a flight. The EMC call occurs to represent these communications. The pilot then has the option to respond at any time during the flight in the EMC call EMCAction.

Vars Sent to ACT-R

CommType – The type of communication message that was sent (e.g. Runway to land on from ATC, Chatter between aircraft, etc)
EMCFlag – Flag representing which EMC call is currently being made

Vars Returned to IMPRINT

None – No variables are returned to IMPRINT, the approach taken is that in a real situation, the pilot may receive the communication right away but may not respond immediately. Pilot response is taken into account in the EMCAction call. This way there is the opportunity for a pilot to prioritize when to respond.

EMCInit

Purpose: To initialize ACT-R with the beginning settings of the IMPRINT model. This EMC works in conjunction with another EMC call named EMCInitUpd. This EMC call and the EMC EMCInitUpd are made at the very beginning of model. These calls occur many times and no simulation time will be used. What happens is IMPRINT asks ACT-R what control/display value it needs; ACT-R returns the control/display that it doesn't have an initial value for, and then IMPRINT returns the value of that control/display.

Vars Returned to IMPRINT

ActionCon – The control/display that ACT-R needs the initial value for
InitEMC – A flag indicating ACT-R has all the initial display/control values it needs

Vars sent to ACT-R

EMCFlag – Flag representing which EMC call is currently being made

EMCInitUpd

Purpose: To initialize ACT-R with the beginning settings of the IMPRIN model. The EMC returns the initial values to ACT-R the control/displays that ACT-R needs.

Vars Sent to ACT-R

CurDisplay – The current control/display that ACT-R asked for
DisplayVal – The value of the control/display
EMCFlag – Flag representing which EMC call is currently being made

Vars Returned to IMPRINT

None – IMPRINT assumes all variables and there values were successfully sent to ACT-R and do not need any return variables to continue.

APPENDIX B
ACT-R COMMUNICATION PROTOCOLS

While IMPRINT and ACT-R are both simulations, they use different representations for the world. Within IMPRINT, all state information is represented as numbers or arrays of numbers. These numbers are stored using global variables. The general way to send or receive data from IMPRINT is by using a COM feature built into the system.

ACT-R's basic representation is memory chunks and production rules. However, since it is built in LISP a sophisticated user can interface with the system in a variety of ways. By compiling the system in Allegro Common LISP we used the provided OLE libraries to interface IMPRINT through with Microsoft's COM.

COM Layer

When IMPRINT starts to run a model it launches and connects to an application. This is done using COM's Client-Server protocol [Box, 1998]. Before an application can be launched by IMPRINT, it must be recorded in Microsoft Window's Registry. This is done by first opening the ACT-R application and executing the LIA function (register-server). This needs only to be done once for each installed platform. When the application is started LIA will automatically load the file "sys:lia-com.fasl". This file contains the necessary COM functions. This file will start a COM server into which IMPRINT can link.

The name of the application and server name is defined by IMPRINT. For the current NASA HPM project, the name chosen by MA&D is "EMC_NASAHPM". With ACT-R, the variable Application-Name is set to "EMC_NASAHPM". The COM server name is constructed from the application name.

IMPRINT has implemented several COM methods. These methods allow an outside application to communicate with IMPRINT. The eight that send and receive data are currently used in LIA. The relevant COM methods are:

(4) ReceiveIntVariable (Name: String; Value: Integer)
(5) ReceiveFloatVariable (Name: String; Value: Float)
(6) ReceiveIntArrayVariable (Name: String; Value: Integer; Index: String)
(7) ReceiveFloatArrayVariable (Name: String; Value: Float; Index: String)
(8) SetIntVariable (Name: String; Value: Integer)
(9) SetFloatVariable (Name: String; Value: Float)
(10) SetIntArrayVariable (Name: String; Value: Integer; Index: String)
(11) SetFloatArrayVariable (Name: String; Value: Float; Index: String)

The first four are used to send data from IMPRINT to ACT-R while the remaining four are used to send data from ACT-R to IMPRINT. IMPRINT sends the variable name and value to ACT-R. The Name is the print name of an IMPRINT global variable. The Value is the value of the variable while the Index is the index into the array.

Even if the data is stored within an array, the data is sent one element at a time. For the Array methods, the index is a string of the form "I J K". For a one-dimensional array the index would be "I 0 0".

The IMPRINT modeler specifies the name of the receiving application. The modeler then decides when and which variables to transmit over the COM link. Incoming data will transparently update IMPRINT's variables.

However, ACT-R has a different point of view. Instead of all data being numbers or array of numbers, ACT-R uses symbols and collections of objects with slots that hold the data. More on how LIA assist in the translations of view points will be discussed later.

External Model Call (EMC) Layer

On top of the COM layer MA&D has developed the EMC protocol. An EMC is a type of remote procedure call (RPC). An IMPRINT integer variable is used to communicate the remote procedure. In the NASA HPM domain, this variable is "EMC_Flag". The LIA variable *Emc-Name* is set to the appropriate string. For any EMC, IMPRINT first sets the EMC_flag to an integer that corresponds to the remote procedure call. It then sends the arguments using the four RECEIVE COM methods. It then sends the value of EMC_flag using the ReceiveIntVariable method. Any return arguments are transmitted to IMPRINT by the remote procedure using the SET COM methods.

The remote application (ACT-R) must first store the data from the COM methods. For any array, this would be one COM method for each element. When the ReceiveIntVariable is invoked with the Name parameter equal to Emc-Name, ACT-R must lookup the corresponding function and apply it to the appropriate arguments. Any values returned by the function, must be associated with their corresponding IMPRINT variable names and sent back using the correct COM methods.

Unlike shared Variables, most of the joint IMPRINT/ACT-R simulations use less than ten EMCs. Since LIA is implemented in LISP, we created a Defemc macro. The macro is similar to the Common LISP Defun macro, which is used to define functions in LISP. Besides defining the top level of an EMC function, the Defemc defines the argument passing and codes needed to dispatch an EMC.

```
(DefEMC (name code return-attributes*) ((var attribute)*) {code}*)

(defemc (emcstart 1 |ACT-RVersNum|) ((version |VersionNum|))
  (assert (= version *IMPRINT-Model-version*) () "Version does not match")
  (lia-reset)
  (if (zerop *trial*) (initiaize-parameters (nth *subject parameters*)))
  ;; Speed Altitude Distance Flap
  (setq *initialize* '(speed altitude flaps)
      *initial-values* nil)
 *ACT-R-Model-version*)
```

In the previous example "emcstart" is the name of a LISP function that is dispatched when the EMC code is 1. This function takes one parameter "version" that is set from the attribute "VersionNum". The value returned is sent to IMPRINT via the attribute "ACT-RVersNum". The body is the LISP code, which executes this EMC.

Attributes

LIA needs to store and decode values. After creating the low level COM link in LISP the desire to generalize the use of IMPRINT variable names as data labels and the mapping of numbers into symbols was the impetus for LIA.

For each IMPRINT variable that is used as a COM label an Attribute is defined. An Attribute is a data structure, which stores the data and tracks the parameters needed for data translation and transmission. The slots within an attribute include:

(4) Name, the IMPRINT variable print name.

(5) Value, the latest decoded value or if the IMPRINT variable is an array then a LISP array of decoded values.
(6) Type, the standard types are Integer, Real, Boolean, and Enumeration. Other domain specific types can be used.
(7) Dimensions, a list of zero to three integers representing the dimensionality of the value. If the list is nil then the data is directly stored in the value slot. Otherwise, an array with the specified dimensionality is stored in the value slot.
(8) Fill, if the value is an array and not all of the array values are initialized then the fill value is used to pad out the array.

The LIA function Find-Attribute takes either a string or a symbol and returns the attribute with the corresponding name. If its input is an attribute then it just returns the attribute.

Besides the normal slot accessors, the setfable function Attribute is defined. Its parameters are (name &rest indexes). The attribute is looked up using find-attribute. If the indexes are nil then its value is returned. If there are indexes then the specified array element is returned or set.

The function Send-Attribute takes an attribute and transmits its value to IMPRINT using the appropriate SetVariable COM method. Before the attribute is sent, its value is encoded as an integer or float. If the attribute is an array, all the elements are sent.

The function Update&Send-Attributes takes a list of attributes and a list of values. First, the attributes are updated using the corresponding values. Then the function Send-Attribute is mapped over the list of attributes. The function Update-Attribute deals appropriately with arrays.

When a ReceiveVariable COM method is invoked, it stores the decoded value into the attribute. The decoding and encoding of values use the generic functions Encode-Value and Decode-Value. The parameters for both of these generic functions are value and type. Type is a symbol for which specialized methods are written. For example:

(defmethod encode-value (value (type (eql ‘boolean))) (if value 1 0))
defmethod decode-value (code (type (eql ‘boolean))) (not (zerop code)))

By writing the two defmethods a new LIA data type can be added. In the NASA domain there existed IMPRINT variables whose data type depended on which display was being examined. Therefore, the data type “Display” was added by simply adding:

(defmethod decode-value (code (type (eql ‘display)))
(decode-value code (enum-bridge (attribute ‘curdisplay))))
(defmethod encode-value (value (type (eql ‘display)))
(float (encode-value value (enum-bridge (attribute ‘curdisplay))) 0.0))

to the glue.lisp file for the NASA code. These methods look up the data type of the attribute and recursively call themselves with that particular data type.

Since the number of variables is large in a model, the attributes are defined in an Excel sheet. This is converted into tab-delimited file and loaded using Load-Table.

Load-Table (file loader layout &key (header 1) (delimiter #/tab) (trim whitespace))

File is the name of the file to be loaded. Loader is a function that processes a row of the table. Layout is a list of column numbers. The data corresponding to the layout list is passed to the loader function. The keyword argument header tells the load-table function how many rows to skip before processing (1 by default). Delimiter is the character that separates the columns of the table (tab by default). Trim is a sequence of characters that are removed for each field (white spaces by default).

The code in the file nasa/code/glue.lisp that loads the attributes is:

(clear-attributes)
(load-table "variables.txt" 'load-attribute '(0 1 5))

Load-attribute takes Name, Type, and Dimensions as parameters. It creates and stores in the attribute table the corresponding attribute. Note the layout '(0 1 5). The Excel sheet contains additional columns that are used by MA&D as documentation.

Enumerations

While IMPRINT uses numbers, ACT-R models use LISP symbols. For the NASA domain, IMPRINT uses codes to represent various discrete items. For example, Flaps are code 600, while the views out the cockpit window are codes 300 – 307. When ACT-R set the flaps the rule was written as *(isa action control flap dial =setting)*. The LIA layer would transparently translate the symbol flap to the integer 600. Similarly when ACT-R rules looked out the window to see if the runway is visable the symbol outwindow-runway would be encoded as 305. To simplify the domain code needed to link IMPRINT and ACT-R the conversions of these domain codes to symbols was integrated into the LIA layer.

The LIA layer calls the mapping of codes to symbols enumerations. Since the set of enumerations is domain specific and numbers almost a hundred for the NASA HPM domain, an Excel file was created to initialize the enumerations table. Both the attributes and enumerations are transmitted via email using an Excel file. The enumeration sheet is also saved as a tab delimited file. It can then be loaded into ACT-R using the form:

(Load-Table "enumerations.txt" 'load-enumeration '(0 1)).

Another standard data type supported by the LIA layer is boolean. For boolean variables the values 1, 0 are mapped to T, nil.

LIA Functions

Synchronization

Since both ACT-R and IMPRINT are simulations, they have to be synchronized. We use a conservative paradigm [Fujimoto, 2000]. One to three EMCs are used to send synchronization events between ACT-R and IMPRINT. In the NASA HPM domain the only synchronization EMC was EMCAction. When these events are sent the attribute "Current" is used as the time stamp for the event. The attribute "ResumeBy" is effectively a "null event" which allows either ACT-R or IMPRINT to advance its internal simulation clock. A null event as defined by Distributed Simulation Systems (DSS) only passes a timestamp. It is a guarantee that the sending system will not produce any messages (events) that will need to be processed until the clock reaches the timestamp.

LIA has several functions that are useful for implementing the synchronization scheme. The first is LIA-Run. This function takes two arguments Current and ResumeBy. It will run ACT-R until either ACT-R halts or the agreed upon time is reached and control needs to return to IMPRINT. LIA-Run returns new values for Current and ResumeBy. These values are to be returned to IMPRINT to complete the synchronization handshake.

LIA-Run can be controlled with several global variables. The variable LIA-Quantum is used as a temporal lookahead in the synchronization paradigm. If the ACT-R rules set the variable *LIA-Waitp* or if there is no goal then the ACT-R clock will be set ahead to the next IMPRINT event. Quitting-Time is used to set a maximum time for the simulation to complete.

When an IMPRINT model that requires communication with an ACT-R model begins executing, an initial communication must take place between ACT-R and IMPRINT. This is done to ensure that both programs are synched up in all areas of the model. This allows ACT-R to reset its current data as well as check to make sure that both IMPRINT and ACT-R are running the same version of the model. In the protocol, when the initial communication is made, the version numbers of both IMPRINT and ACT-R are exchanged; if there is a mismatch then the model will end.

The function LIA-Reset is used to initialize a new run in a multi-trial simulation. It resets ACT-R's model, assigns an offset to ACT-R's clock and tracks which run and trial the simulation is performing. This function is usually called by EMCStart.

The third generic protocol is a final communication between IMPRINT and ACT-R that allows both programs to end correctly. The final event in the IMPRINT model is a communication between IMPRINT and ACT-R where IMPRINT sends over information to let ACT-R know that the IMPRINT simulation is ending and that this will be the final communication between the two programs. In this protocol, IMPRINT will send over a flag to ACT-R. ACT-R will then acknowledge the end of the simulation. IMPRINT will then end its simulation.

The function Signal-Done can be used in an ACT-R's model to end the simulation and set a completion code. While the function, LIA-WrapUp is used to clean up any pending goals in ACT-R. LIA-WrapUp is usually called by EMCEnd which should be the last EMC sent from IMPRINT to ACT-R.

After running the simulation, IMPRINT will disconnect from the COM server started by ACT-R. If the LIA variables *auto-exit* is true then the ACT-R application will terminate itself in about 3 seconds after the disconnect.

Attention-Situation Awareness (A-SA) Model

Chris Wickens, Jason McCarley, and Lisa Thomas
University of Illinois at Urbana Champaign
Urbana, Illinois

A. Description of Modeling Effort

A1. Foundation of the model.

The underlying theoretical structure of the Attention-Situation Awareness (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. The first module corresponds roughly to Endsley's (1995) Stage 1 situation awareness, the second corresponds to her Stages 2 and 3. In dynamic systems, there is a fuzzy boundary between Stage 2 (understanding) and Stage 3 (prediction) because the understanding of the present usually has direct implications for the future, and both are equally relevant for the task.

The elements underlying the attention module are contained in the **SEEV** model of attention allocation, developed by Wickens, Helleberg, Goh, Xu, and Horrey (2001; Wickens, Goh, Horrey, Helleberg, & Talleur, 2003), and are shown schematically in Figure 1 (McCarley, Wickens, Goh, & Horrey, 2002). These elements indicate that the allocation of attention in dynamic environments is driven by bottom up attention capture of **salient** events, is inhibited by the **effort** required to move attention (as well as the effort imposed by concurrent cognitive activity), and is also driven by the **expectancy** of seeing **valuable** events at certain locations in the environment. The first letter of each of the four boldfaced terms, defines the SEEV model.

In Wickens and McCarley (2001), we applied a version of this attention model, coupled with a version of an inference model based on the belief updating model of Hogarth and Einhorn (1992), to develop a version of the A-SA model that could be applied to predicting errors in taxiway navigation, as shown in Figure 2. In this modeling effort, greatest emphasis was placed on carefully defining parameters of salience of events, the effort (workload) of ongoing activity and dividing attention between multiple events, and on the **relevance** of events. Relevance was assumed to correspond to the value of those events to the pilot's task of understanding, or maintaining awareness of which direction he was supposed to turn at a given taxiway intersection. This understanding was modeled by the belief updating module with memory decay, and ranged from perfect (in which case a turn was always correct) to absent, in which

SEEV MODEL

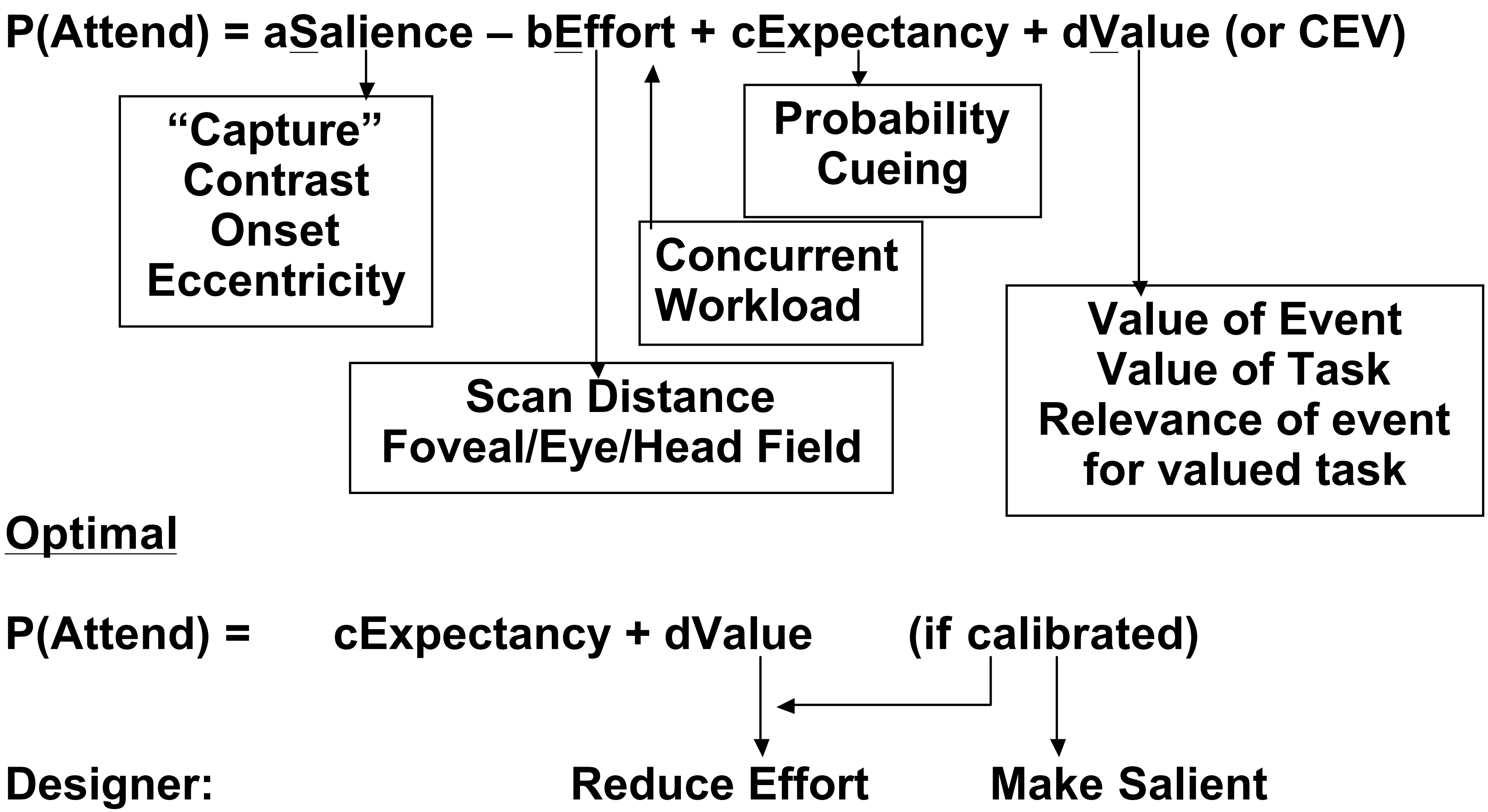


Figure 1. The SEEV model.

EVENTS
E(C,V): Conspicuity, Info Value (relevance to situation of interest)

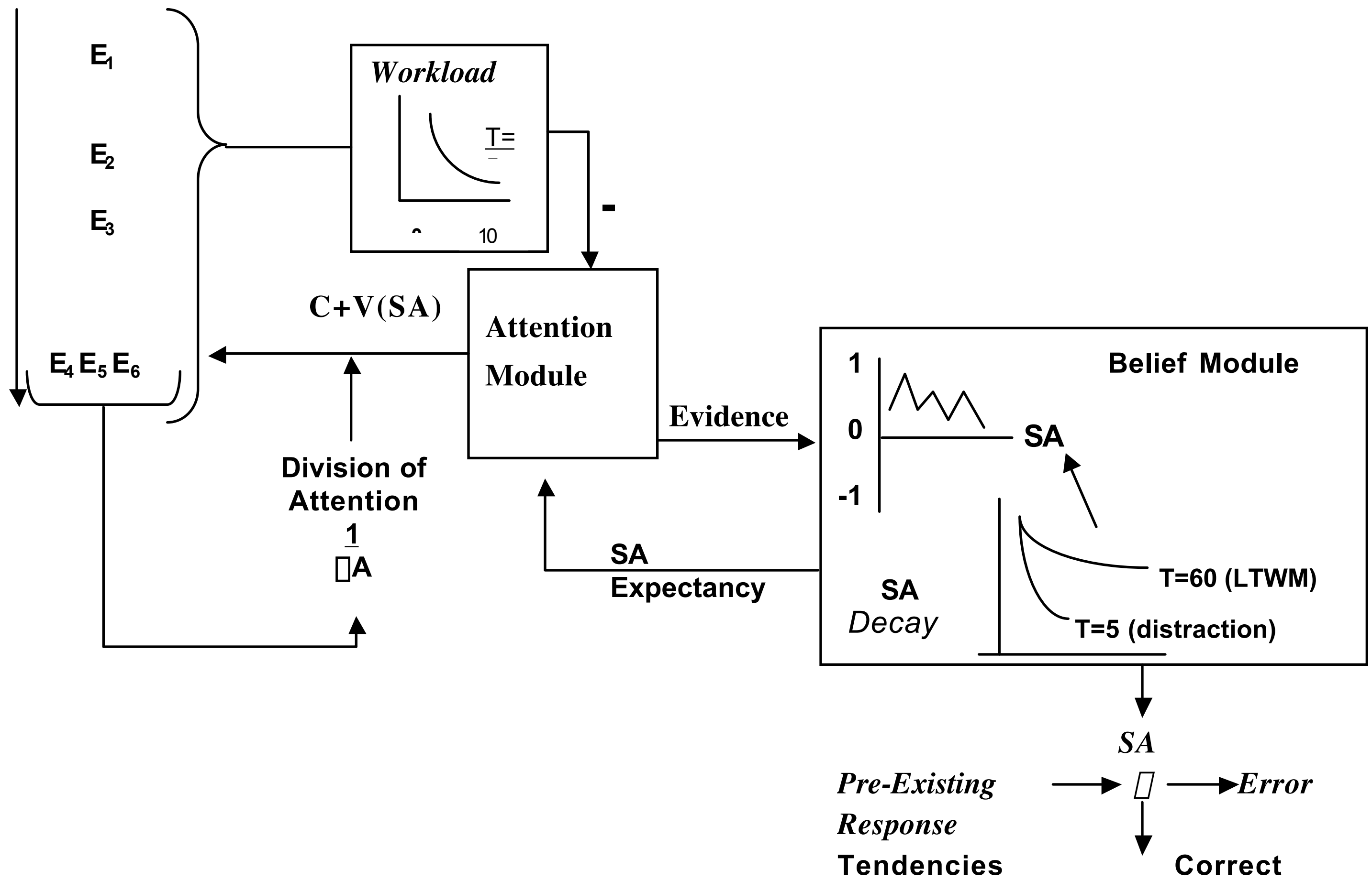


Figure 2. A-SA Model for taxi-way error prediction.

case a turn was governed by pre-existing response tendencies, which could therefore trigger a wrong turn error. We were provided by NASA with a rich set of taxiway turn errors for model validation performance data, and we were directly able to attribute these errors to a loss of situation awareness.

In the second year of the project, we were asked to apply the model to a very different sort of data, describing pilots performing simulated approaches to an airport, when supported or not supported by a synthetic vision system (SVS) display, intended by designers to support situation awareness. The NASA part-task simulation and data set supplied to us are described elsewhere in this proceedings (Goodman, Hooey, Foyle & Wilson, 2003). Several things about this new validation effort required us to modify our modeling approach from that used in the first year. First, loss-of-SA incidents were now quite scarce in the data provided by NASA. Second, we did not have available any explicit or implicit "probes" of SA (e.g., SAGAT measures; Endsley, 1995) that might also have availed data for modeling. Third, although we were provided with a full set of data records in both video and digital files, these revealed few discrete "events" that could be tied to the gain or loss of SA, in the same manner that the events from the taxiway data had been able to do. With fewer "events" it became more difficult to employ the salience component of the SEEV model, since salience serves the model only to the extent that it can be defined as a direct property of a discrete event.

To compensate for these shortcomings of the current data set, we were provided an extensive set of eye-movement data, which, in contrast to the first year taxi-data, we could now model directly as the output of our attention module. In addition, while we did not have events defined by salience, we **did** now have available channels defined by distinct locations. Following the precedence of our previous scanning model approaches, we define these channels as **Areas of Interest (AOI)**. Each AOI can be defined in terms of a (1) **transition to it**, or "**visit"** (from another AOI), a (2) **dwell duration** on the AOI before leaving it, and a (3) **percentage dwell time** looking at it (which is the product of the frequency of visits and the mean dwell duration, divided by the total amount of time). While we could not thereby model the salience of events, we were able to model the **effort** of moving attention (transitioning) from one AOI to another, assuming that such effort is monotonically related to the distance between AOIs. These distances could be derived from the simulation dimensions provided to us by NASA. Furthermore, since the approach/landing task is one that has been often studied within the aviation domain, we were able to define the **value** of tasks on the well established hierarchy of **aviate > navigate**.

Following the procedures developed in Wickens, Helleberg, Goh, Xu, and Horrey (2001; Wickens et al., in press), we modeled the value of an AOI to be the value of the task served by the AOI multiplied by the relevance of that AOI to the task in question. Finally, also following similar procedures to those used in Wickens et al. (2001), we modeled the **expectancy** for information contained in an AOI, in terms of the **bandwidth** of information in that AOI (that is, the frequency with which events or changes occurred to information contained within the AOI). Thus we were able to estimate the quantitative parameters necessary to predict how frequently an AOI optimally **should** be visited, as a function of the momentary relevance and bandwidth of that AOI, and to predict how frequently it **will** be visited, given the inhibiting influence of effort (which inhibits scans over wide visual angles).

A2. General Approach to Modeling

Figure 3 provides our schematic representation of the approach to the landing used in the current SVS simulation. Importantly, each approach in the 10 scenarios that were described by NASA can be subdivided into four phases, distinguished from each other by potential changes in relevance and bandwidth (in some scenarios):

- Phase 1. Above 1000 ft. Regular "steady state" flight.
- Phase 2. 1000 ft – 850 feet. Lined up on runway (whether visible or not).
- Phase 3. 850-600 feet. Runway becomes visible as the airplane drops below the cloud ceiling in most IMC scenario landings.
- Phase 4. Below 600 feet. Runway remains hidden in low-visibility go-around scenarios.

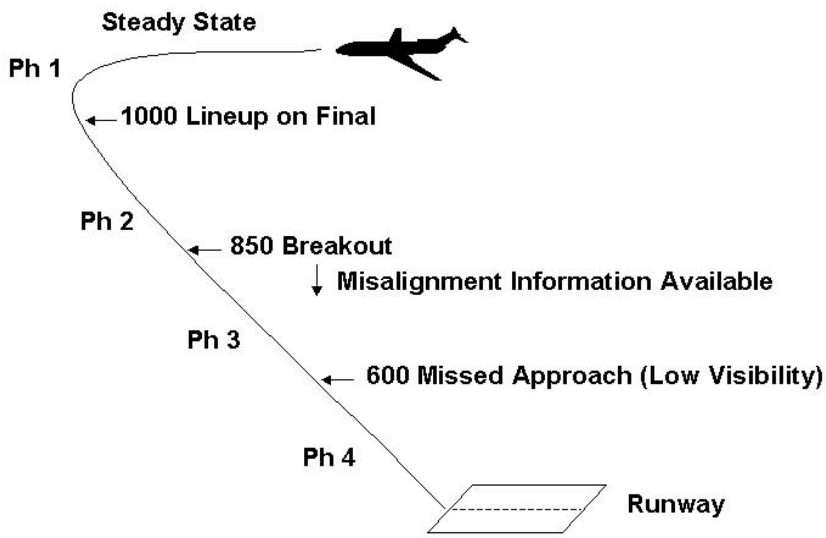


Figure 3: Schematic representation of scenario time line

Each of these phases defines a separate eye-movement data base to be analyzed. With this representation of the data, we applied four different approaches to the analysis, as shown in the 2x2 matrix of Figure 4. The figure differentiates the extent to which we are interested in the common general behavior of all pilots (left) or differences in the specific behavior of individual pilots (right), and the extent to which our modeling efforts are applied to pilot performance (top row) versus applied to visual scanning (bottom row). While the general scan data were modeled for all six scenarios (5-10) that were flown in IMC, we chose to model in detail, two landing scenarios provided by NASA, because both were characterized by some performance data, from which variability in situation awareness (between pilots) could be inferred. These were scenario 6, a baseline scenario flown in IMC, in which a mismatch between the visible runway and the ILS instrument forced a go-around below 850 feet, and scenario 10, in which the same mismatch was reflected in a misalignment between the SVS display, and the runway view.

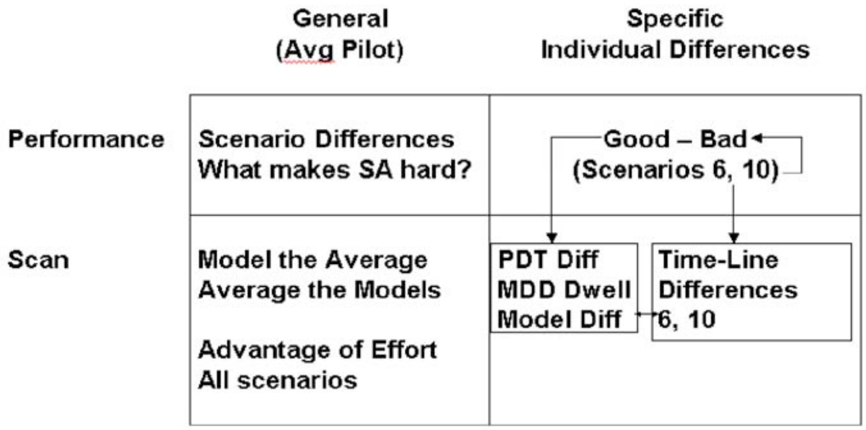


Figure 4: Different approaches to data analysis

In the upper right cell of Figure 4, the quality of situation awareness was operationally defined from performance measures by the speed with which pilots became aware of the misalignment in the two scenarios. Careful review of the video tapes and transcriptions revealed that in both scenarios, pilot 5 maintained good SA, rapidly noticing the misalignment and executing the missed approach, whereas pilots 3 and 4 either noticed this only after a considerable delay, or not at all, needing to be reminded by the confederate first officer. The distinction between the two "classes" of pilot behavior ("good" and "bad" SA) was important, allowing us to discriminate their attention allocation behavior, as we describe below

A3. Implementation of the Model

In implementing the model, we estimated the coefficients of bandwidth, relevance and task priority, as shown in Figure 5a (scenario 6) and 5b (scenario 10) both involving a missed approach. Using the same procedures applied by Wickens et al. (2001, 2003), we employed the "least integer ordinal value" heuristic, in assigning these coefficients. This is a heuristic that maintains integer values of all coefficients, and tries to keep these as low (and therefore simple) as possible, while preserving any necessary ordinal relations. As one example, the bandwidth of the instrument panel during a missed approach (climbing, accelerating, turning), is higher than during a straight in approach (compare the bandwidth parameters in Figure 5 above – straight in—and below – missed approach – the 650 ft level). The reader can also see in Figure 5a, that in IMC conditions, when the outside world (OW) is not visible, its bandwidth is assigned a value of 0. As another example, the relevance of both aviating and navigating is increased during the missed approach phase below 650 ft because of the higher criticality of both of these tasks at unusual attitudes and low to the ground.

In addition to these parameters, integer parameters for effort were assigned on the basis of the distance between displays, as extracted from the documentation provided by NASA (**Figure 2 in Goodman, Hooey, Foyle & Wilson, 2003**). Laterally adjacent displays imposed an effort value of 1, vertically separated displays, an effort value of 2, and the presence of an intervening displays also imposed an effort value of 2 for scans between the two flanking displays.

S10

Parameter		Above 1000 ft	1000-800	800-650	Below 650
Bandwidth (B)	IP	3	3	3	5
	OW	0	0	2	3
	ND	1	1	1	1
	SVS	2	2	2	3
Relevance (R)	IP (av)	2	2	2	3
	IP (nav)	1	1	1	2
	OW (av)	0	0	1	0.5
	OW (nav)	0	0	2	1.5
	ND (av)	0	0	0	0
	ND (nav)	2	1	1	2
	SVS (av)	1	1	2.5	0.5
	SVS (nav)	1	2	2	1.5
Priority (V)	Aviate	2	2	2	4
	Navigate	1	2	2	3

Figure 5b: Coefficient Values for Scenario 10

Our model calculated the "attractiveness" of an AOI as a direct function of the product of its bandwidth and relevance (the latter, modulated by the value of the task to which AOI it was relevant). To this product was added three *minus* the effort required to reach that channel from the currently-attended AOI. This ensured that an increase in the effort needed to reach an AOI would decrease the attentional weight of that AOI. Thus, when effort needed to reach an AOI was minimal (a value of 0) , attentional weight of the object was increased by 3. When the effort needed to reach an AOI was maximal (a value of 3), attentional weight of the object was not increased relative to the value established based on bandwidth and relevance. The probability with which attention shifted to a given AOI was then calculated by dividing the attentional weight of that AOI by the summed attentional weights of all AOIs. The model thus predicted the movement of attention from one AOI to another.

We developed two versions of the model. In version 1, upon attention landing on an AOI, the relevance of that AOI became zero (reflecting the acquisition of needed dynamic information from that AOI), and the attentional weights of the remaining AOIs were recalculated so that the next gaze could occur. In version 2, the relevance would remain the same after a fixation, and the model would either move to another AOI **or** remain on the AOI with a probability that was related to the relative attentional weights of all available AOIs (including that of the momentary fixation). In this way, the eye could remain for a **long dwell** on a particular AOI, of high relevance.

Thus, for each phase of flight (characterized by its unique set of values as shown for two of the scenarios in Figure 5), the model generated an NxN transition probability table, where N was set equal to the number of "active" AOIs. For example, as can be seen in Figure 5a, in scenario 6 (baseline), N was equal to 3 (Instrument panel, IP, navigational display ND, and Outside World, OW), whereas in Figure 5b (SVS) N was equal to 4, since the SVS display itself defined a now-active AOI. Table 1 presents the active AOIs for scenarios 5 through 10, in each of the four phases. (Fixations on the MCP or other areas were sufficiently rare that these were not included in our modeling efforts).

The cell values of a transition table from model version 1 defined the predicted probability of t**ransitioning** from one AOI to another. The row (or column) values within the table provided the predicted **number of visits,** or "popularity" of an AOI. In version 1, if we assumed that the number of visits was correlated with the length of each visit, we could label this as a prediction of the percentage dwell time (PDT) on each AOI. In model version 2, we were able to predict the PDT directly by summing the total time of dwells on each AOI, and dividing by total trial time. In either case, we were then able to correlate these two model-derived measures (transitions within each cell and PDT), with the actual scanning behavior extracted from the eye movement records of each pilot, in each of the 4 phases of each scenario, to provide measures of model validation. Finally, it was possible to compute such correlations in either of two ways: (1) compute the correlation of each pilot individually, and average the correlation, (2) compute the average scanning data across the 3 pilots, and correlate these averaged data with the model prediction (only two pilots provided scanning data for scenario 6).

Table 1. Active AOI's (✓) for each scenario and phase.

Phase →

Scenario	**AOI**	**1**	**2**	**3**	**4**
5	OW				
	IP	✓	✓	✓	✓
	ND	✓	✓	✓	✓
	SV				
6	OW			✓	✓
	IP	✓	✓	✓	✓
	ND	✓	✓	✓	✓
	SV				
7	OW			✓	✓
	IP	✓	✓	✓	✓
	ND	✓	✓	✓	✓
	SV	✓	✓	✓	✓
8	OW			✓	✓
	IP	✓	✓	✓	✓
	ND	✓	✓	✓	✓
	SV	✓	✓	✓	✓
9	OW			✓	✓
	IP	✓	✓	✓	✓
	ND	✓	✓	✓	✓
	SV	✓	✓	✓	✓
10	OW			✓	✓
	IP	✓	✓	✓	✓
	ND	✓	✓	✓	✓
	SV	✓	✓	✓	✓

B. Findings

B1. Modeling the average pilot.

Using model version 1, we computed the model fitting correlations based on transitions, and based on 4 different versions of the PDT predictions. These four versions involved the presence of absence of the **effort** parameter, and involved either correlating model predictions with the average scanning behavior of the 3 pilots, or averaging the correlation values {r} across the three pilots. From this exercise with model version 1, we drew the following three conclusions:

1. Methodologically, the correlation with the average scan data (average r = 0.79) is greater than the average of the pilot correlations (r = 0.60).

2. The correlations (model fits) with the PDT (average = 0.60) are generally higher than those with the Transitions (average = 0.42). The reason for the decrease in the latter values is, presumably, that the model predicts symmetric transition probabilities between AOIs (e.g., P(OW→ IP) = P(IP→ OW)), and the data may show asymmetries of direction.

3. The inclusion of the Effort parameter offers no benefit to model fitting, and in fact, in some cases, actually reduces the fit. We conclude from this, that pilots were not inhibited from making longer scans, if there was valuable and high-bandwidth information to be obtained at the more distant AOI. Such a conclusion was also consistent with one drawn by Wickens et al. (2001). However a decision was made to retain the effort parameter in our subsequent modeling efforts because of its cognitive plausibility, and because it could be valuable in modeling differences between pilots.

We subsequently decided to proceed by modeling all individual scenarios with model version 2 (predicting percentage dwell time), as this version generally provided higher model fits across the scenarios and phases; and in particular, those phases in which more than 2 AOIs were "active" (see Table 1). Furthermore, such a model did produce dwell time differences that accounted, partially, for the different percentage dwell durations. Appendix A shows the scatter-plots between predicted and obtained PDT, on each page representing the four approach phases for each pilot. Data points for the four AOIs are labeled in the figure, whether these are "active" or not. Each scatter plot generated a correlation, and Table 2 presents the resulting correlations. From the table one can again discern the generally higher correlations for the model of the average pilot, than for the average of individual pilots. More important however, are the following two observations.

4. The average subject model fits, along with the model fits of individual pilots are all generally good, with positive correlations generally in the .60 to .80 range, although there are some exceptions. In particular, scenario 6 shows lower correlations. One possibility is that scenario 6 had only two pilots contributing eye movement data, and one of these (pilot 3) had very poor individual fits on all three phases (see top line of Table 2, Scenario 6). This point will be important below. There does not appear to be any common trend that would discriminate higher correlations from lower ones (i.e., later versus earlier phases, or SVS versus non-SVS scenarios).

Table 2. Correlations between model version 2 (with effort) and obtained data on percentage dwell time. (Eye movement data were missing for subject 4 scenario 6.)

						Description of Scenario
Scen05						
	TOTAL DWELL TIME					
		Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	
	s03	0.738	0.841	0.513	0.776	Missed approach
	s04	0.696	0.721	0.768	0.776	(Low ceiling)
	s05	0.728	0.900	0.655	0.591	Baseline display
	Average Subject	0.756	0.839	0.466	0.813	
	Average of subjects	0.721	0.821	0.645	0.714	
Scen06						
	TOTAL DWELL TIME					
		Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	
	s03	0.119	-0.086	-0.174	0.207	Missed approach
	s04					Terrain (ILS)
	s05	0.611	0.703	0.985		Mismatch
	Average Subject	0.347	0.290	0.457	0.207	Baseline display
	Average of subjects	0.365	0.309	0.406	0.207	
Scen07						
	TOTAL DWELL TIME					
		Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	
	s03	0.747	0.463	0.694	0.813	Normal landing
	s04	0.310	0.960	0.903	0.978	SVS display
	s05	0.640	0.732	-0.302	0.081	
	Average Subject	0.664	0.895	0.724	0.921	
	Average of subjects	0.566	0.718	0.431	0.624	
Scen08						
	TOTAL DWELL TIME					
		Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	
	s03	0.819	0.747	0.194	0.325	Late runway sidestep
	s04	0.652	0.850	0.677	0.409	SVS display
	s05	0.581	0.763	-0.383	0.243	
	Average Subject	0.643	0.645	0.692	0.875	
	Average of subjects	0.684	0.787	0.163	0.326	
Scen09						
	TOTAL DWELL TIME					
		Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	
	s03	0.670	0.605	-0.121	0.346	Missed approach
	s04	0.710	0.921	0.253	0.774	(Low ceiling)
	s05	0.594	0.251	0.816	0.535	SVS display
	Average Subject	0.637	0.751	0.278	0.771	
	Average of subjects	0.658	0.593	0.316	0.552	
Scen10						
	TOTAL DWELL TIME					
		Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	
	s03	0.638	0.858	0.276	0.103	Missed approach
	s04	0.701	0.912	0.536	0.420	(Terrain mismatch)
	s05	0.815	0.347	0.363	0.163	SVS display
	Average Subject	0.664	0.780	0.583	0.898	

5. A review of the scatter plots of the individual subject data (see Appendix) appears to reveal one consistent trend, which would account for a drop in the correlations. That is, often when the correlation is low, it is because OW scanning occurs much more frequently than predicted by the model, whereas rarely if ever is OW scanning done less than predicted. There are two possible explanations for this. First, in phases 1 and 2, the OW is invisible (IMC), and our model therefore predicts no "relevance" for either aviating or navigating tasks (see Figure 5). However it is likely that a vigilant pilot, knowing that visibility will be required for landing, will occasionally glance to the OW to assess whether anything is visible. Second, in phases 3 and 4, when the OW **is** visible (in all but scenarios 5 and 9, which were missed approaches because of the low clouds). The increased OW scan could reflect the fact that the OW was used both for aviating (the true horizon, rather than the IP or SVS horizon) as well as navigating (the true runway, rather than the SVS runway), thereby increasing the "relevance" coefficients (for the OW for aviating) from the model values that we assumed and assigned in Figure 5.

B2. Modeling individual pilot differences.

As we discussed in section A2, there appeared to be differences between pilots in their awareness of the runway offsets in scenarios 6 and 10. Hence we also asked if differences in the model fits might have accounted for the distinct differences between pilot 5 on the one hand, who appeared on the basis of the offset-noticing performance data, to have "good SA", and pilots 3 and 4, who did not because they either failed to notice, or noticed very late, the misalignments. The data provided only modest support for this difference. For scenario 6, phase 2, which would characterize the scan pattern just prior to the information regarding the misalignment becoming available, the model fit for pilot 5 ($r = .70$) is much better than for pilot 3 ($r = -.08$). (There were no eye movement data for pilot 4.) More detailed examination of these differences revealed that pilot 5 looked at the PFD (a scan necessary to notice the misalignment), whereas pilot 3 did not look there during this phase (see Appendix).

For scenario 6, phase 3, similar findings revealed better model fits for pilot 5 ($r = .98$) than for pilot 3 ($r = -0.17$). More detailed examination of the scan records suggests that the lower model fit results because pilot 3 again fails to look at the PFD, where evidence of the misalignment is present.

Finally, for scenario 10, phase 2, the findings are far less consistent. Indeed pilot 5 appears to show a much poorer model fit than do pilots 3 and 4. The cause of the low prediction of pilot 5 for scenario 10 phase 2 appears to be scanning much longer on the SVS than the model predicts and much less on the PFD, an issue that will be addressed below. Similarly, during phase 3, when the discrepancy should be noticeable, there is again a poorer model fit for pilot 5, than for pilot 4 (although 5 is better than 3). Thus, in general, there is not a great deal in the model fitting that predicts differences in scan pattern during the critical time period (phase 3) when the discrepancy should be noticed. To seek these differences we turned to a final level of descriptive analysis of eye movements using a qualitative approach which departed from our quantitative modeling approach.

B3. Describing pilot differences in scanning behavior.

The A-SA model accounts for stable, averaged scanning data, but in its current form does not account for the micro-strategies underlying specific scans or transitions. Yet such micro-strategies appear to underlie, in particular, the difference between the one "good" SA pilot (#5) and the two other pilots (#3 & 4). These micro-strategies are illustrated in Figure 6 which shows the scan pattern and associated dwell durations for phase 3 of scenario 6 (Figure 6a, pilots 3 and 5) and of scenario 10

(Figure 6b, pilots 3, 4 and 5). This phase, (phase 3), is chosen since it is the one during which information regarding the discrepancy between the outside world and the instruments (PFD: S6 or SVS display: S10) becomes visually available. The time in the sequence of the "noticing response" – initiation of a go-around – by pilot 5 is indicated by the * in the figure. This time was extracted from the time line record of flight control parameters provided to us by NASA.

In both scenarios, pilot 5 shows two features: (1) a direct scan between the two relevant AOIs, reflecting, we infer, such a comparison necessary to notice the mismatch; (2) a long dwell (>5 seconds) on the outside world, reflecting, we infer, a careful evaluation of the information content of that AOI (runway location) with the long dwells required to notice the discrepancy. The initiation of the go-around occurs within one second after this fixation pattern is completed. Neither pilot 3 nor 4 show both of these features of the discrepancy-noticing fixation in conjunction, and in particular, neither pilot shows a long dwell on the outside world.

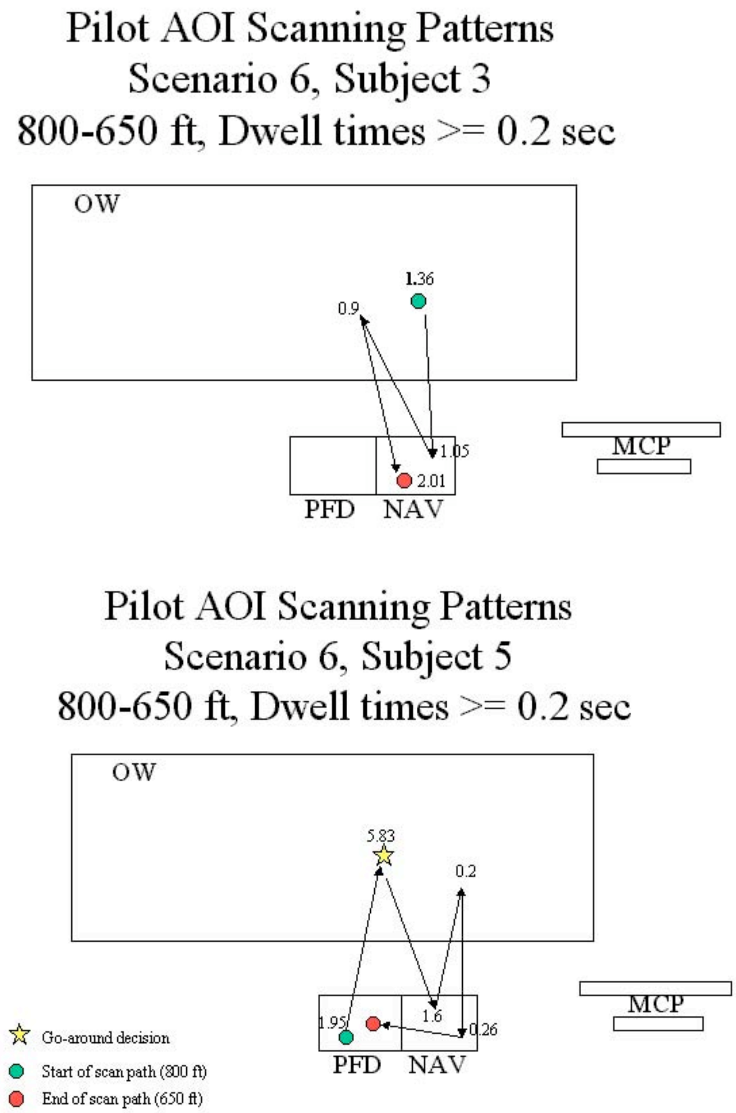


Figure 6a: Qualitative analysis of noticing runway offset, an awareness which was successful for subject 5, but not for subject 4. Scan path is Scenario 6.

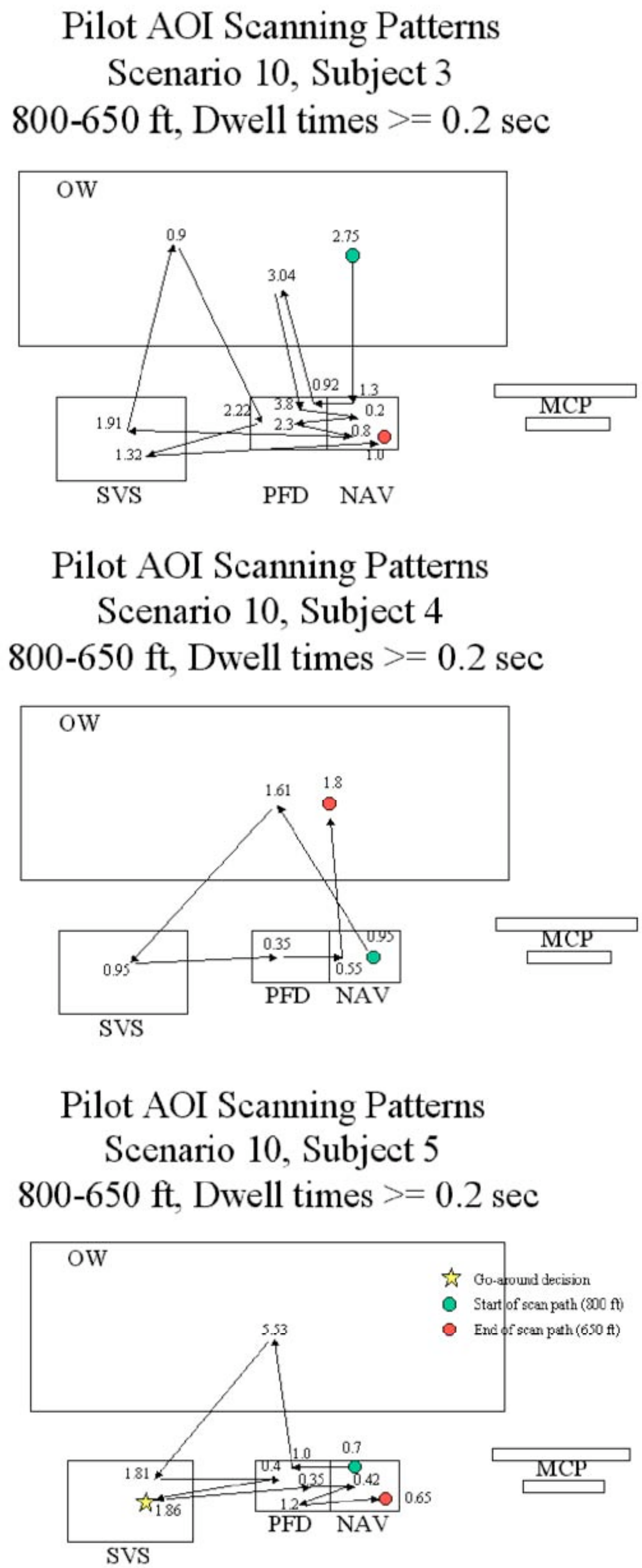


Figure 6b. Qualitative analysis of noticing runway offset, an awareness which was successful for subject 5, but not for subject 4. Scan path is scenarios 10.

C. Implications

One implication of the current modeling exercise appears to be that, in the environment modeled here, the effort of making longer scans does not appear to inhibit those scans. That is, the model fit is just as good, when driven by only bandwidth and relevance, as when effort is included. Such a conclusion is consistent with our findings in previous research (Wickens et al., 2001), that scanning of instrument rated general aviation pilots can be very effectively modeled with only expectancy and value as parameters.

A second implication is inherent in the better model fit of the "good" SA pilot (pilot 5), relative to pilot 3, a discrimination that provides some validation of the model in scenario 6.

A third implication is that the wide individual differences that appear to exist within the data provided, may be modeled as much by the dwell duration (see Figure 6), as by the particular transition. Such a distinction is one drawn by Harris and Christhilf (1980), and Bellenkes, Wickens, and Kramer (1997) between short dwells, designed to confirm hypotheses, and longer dwells, designed to acquire new visual information. However the dividing line between these two forms of dwells here (around 2 seconds) is generally longer than that observed by Bellenkes et al and by Harris and Christhilf (around 1 second) This discrepancy can in part, be accounted for by the fact that those investigators did not examine scanning in off-normal scenarios. It is also appropriate to note that our operational definition of "dwell duration" is not the length of a single fixation at a point location in space, but rather, refers to the duration of a repeated series of consecutive fixations within a wider AOI, before that AOI is left.

D. Lessons Learned

In terms of our modeling effort, we are learning the importance of incorporating dwell duration into our modeling. The initial step in this direction was taken with model version 2, although we have not yet exercised the model to predict the mean dwell duration.

We believe that our particular modeling effort suffered from the paucity of direct "situation awareness" measures available in the current simulation, and hence a "lesson learned" might be the need to obtain explicit SA measures collected in future simulations. That is, as noted, our primary focus has been on modeling Stage 1 SA (Attention and noticing events, inferred from scanning), rather than modeling the objective of Stage 1 SA, inherent in Stages 2 and 3 (understanding and prediction). We were not able to firmly link the former to the latter, because of the paucity of data that could be used to infer the presence or absence of Stages 2 and 3 SA. Correspondingly a more robust test of the model can be achieved with data from a greater number of pilots. This would provide a wider range of responses to off-normal events, a criterion that could be used for performance based model validation. In this context we did feel fortunate that the pronounced differences between the two classes of pilots (5 vs. 3&4) emerged.

References

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. Aviation, Space, and Environmental Medicine, 68(7), 569-579.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1), 85-104.

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.

Harris, R. L., & Christhilf, D. M. (1980). What do pilots see in displays? In G. Corrick, E. Hazeltine, & R. Durst (Eds.), Proceedings of the 24th Annual Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. Cognitive Psychology, 24, 1-55.

McCarley, J. S., Wickens, C. D., Goh, J., & Horrey, W. J. (2002). A computational model of attention/situation awareness. Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors and Ergonomics Society.

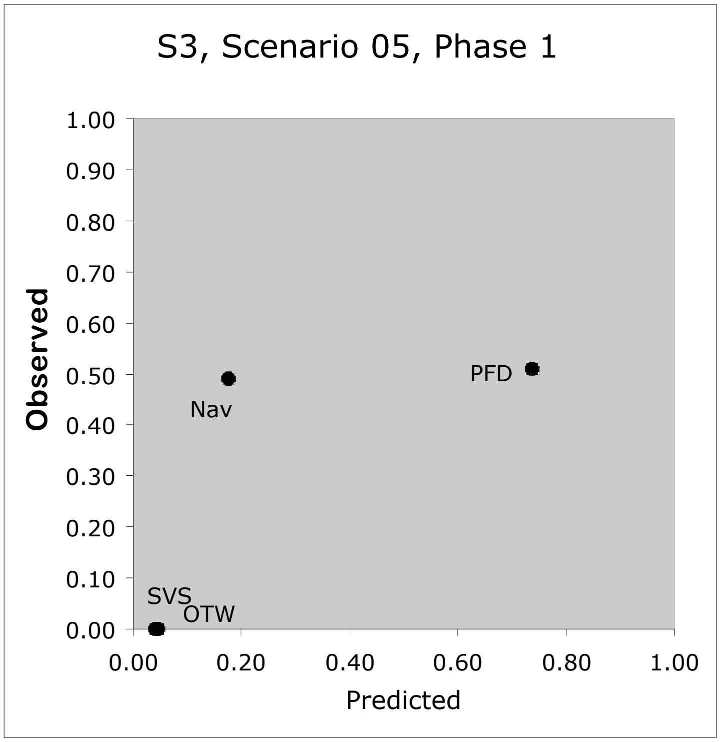
Wickens, C.D., Goh, J., Horrey, W.J., Helleberg, J.H., & Talleur, D.A. (2003). Attentional models of multi-task pilot performance using advanced display technology. Human Factors, 45, #3, 360-380.

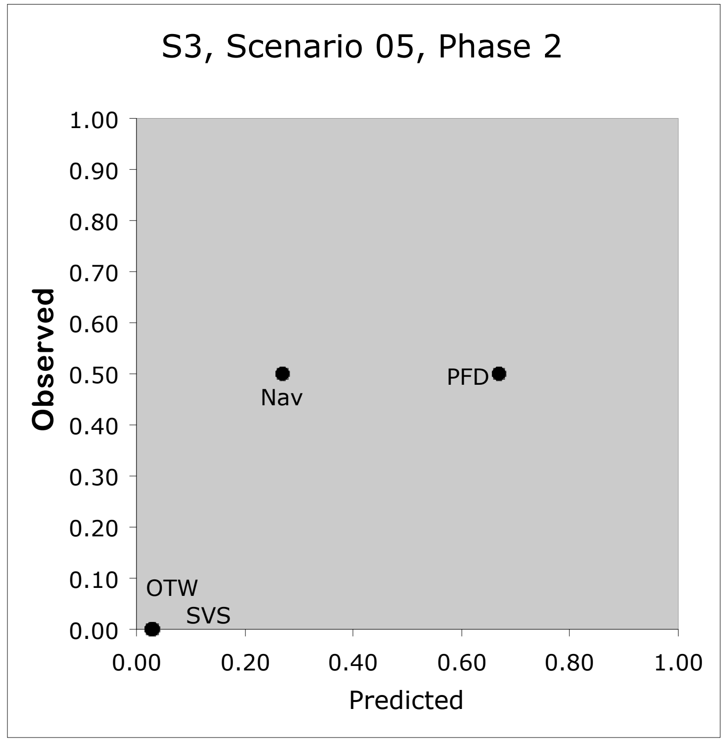
Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, B. (2001). Pilot task management: testing an attentional expected value model of visual scanning (ARL-01-14/NASA-01-7). Savoy, IL: University of Illinois, Aviation Research Lab.

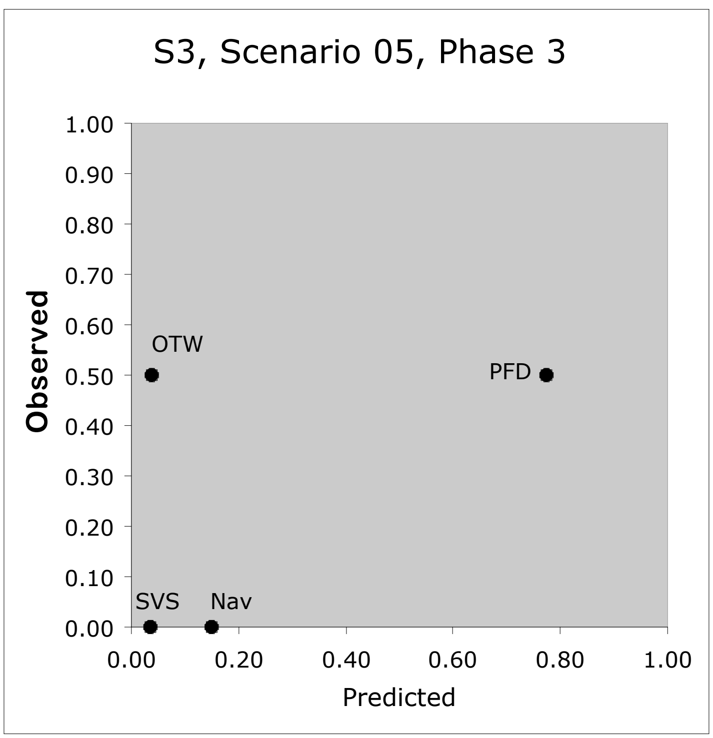
Wickens, C. D., & McCarley, J. S. (2001). Attention-situation awareness (A-SA) model of pilot error (Final Technical Report ARL-01-13/NASA-01-6). Savoy, IL: University of Illinois, Aviation Research Lab.

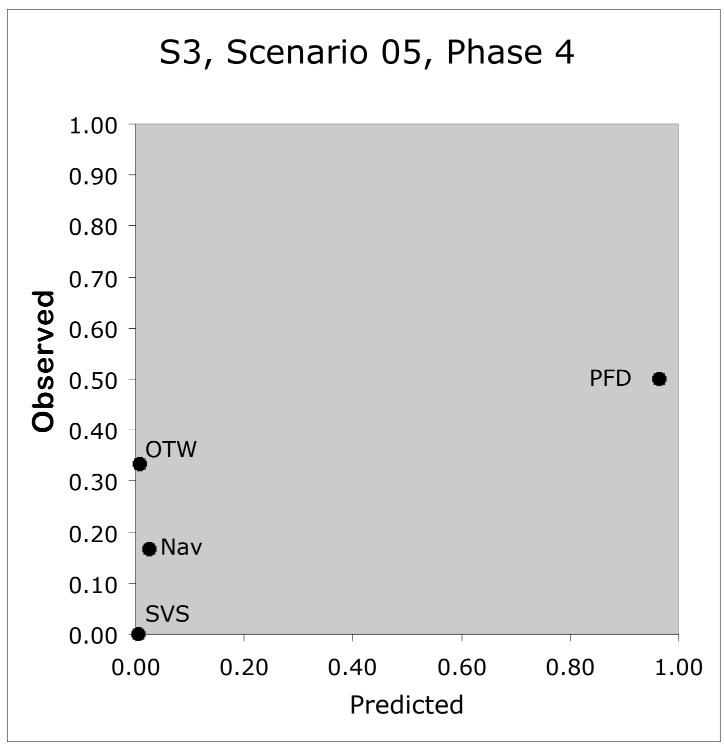
Appendix: Scatter plots from model version 2.

Scenario 5, Subject 3, Phases 1-4

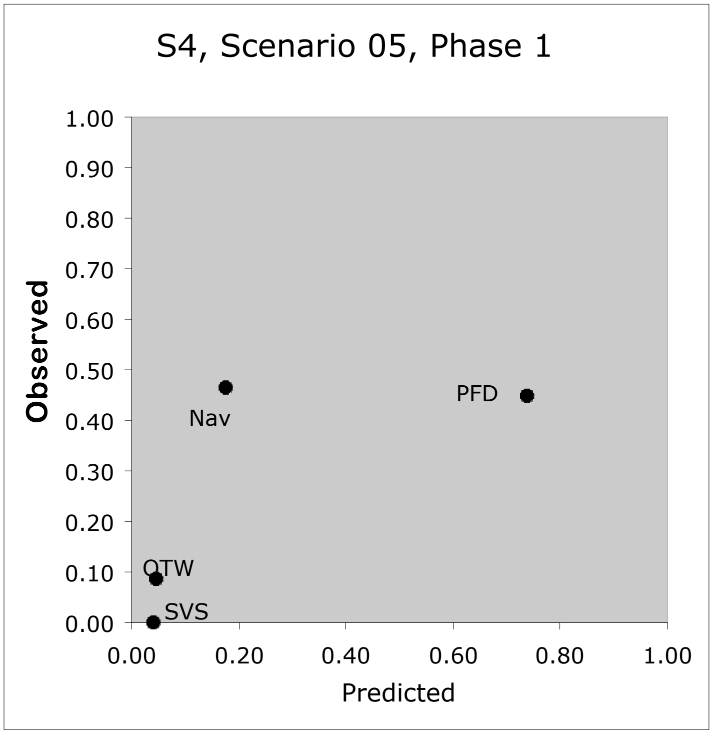


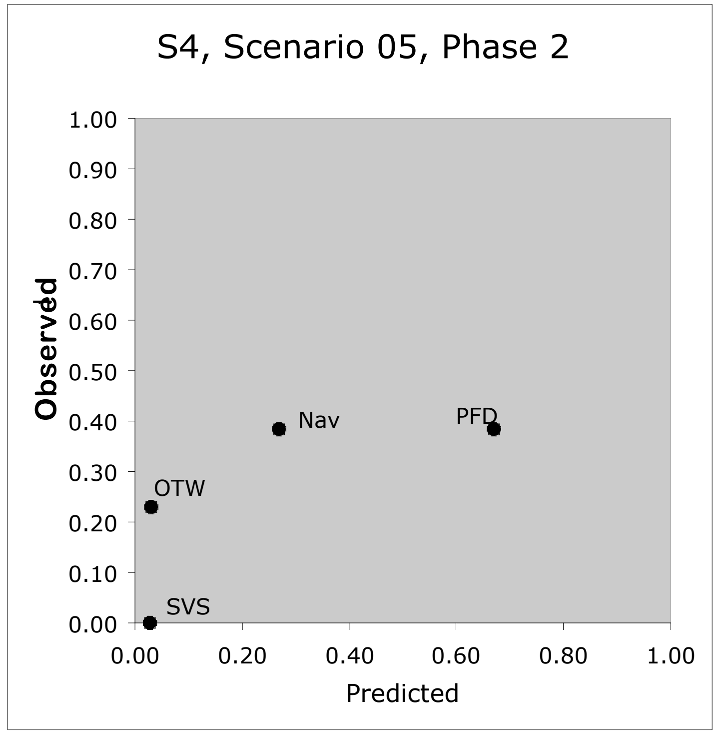


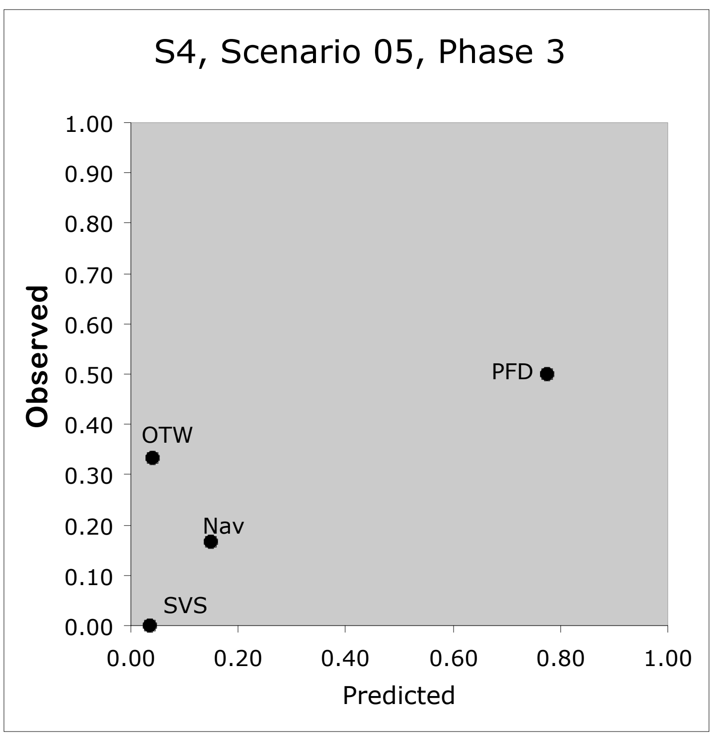


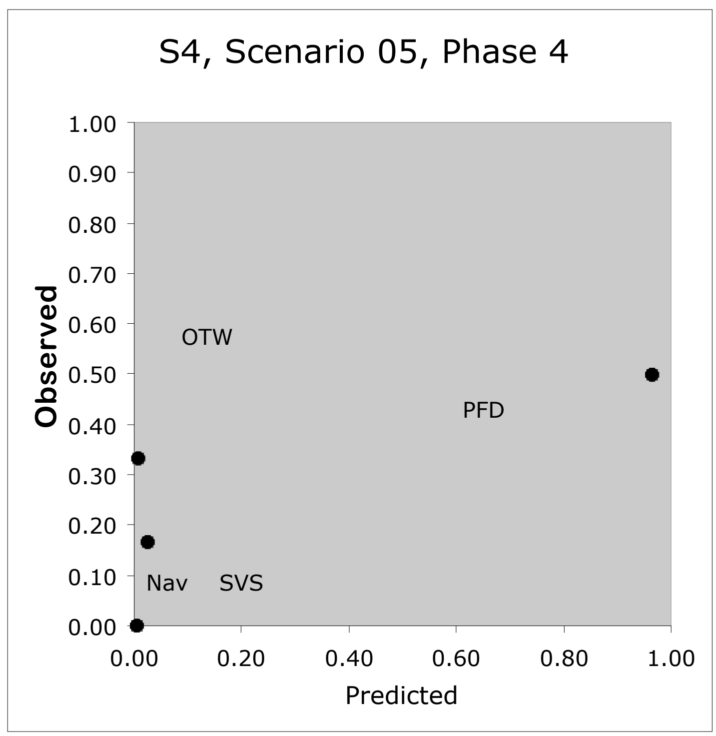


Scenario 5, Subject 4, Phases 1-4

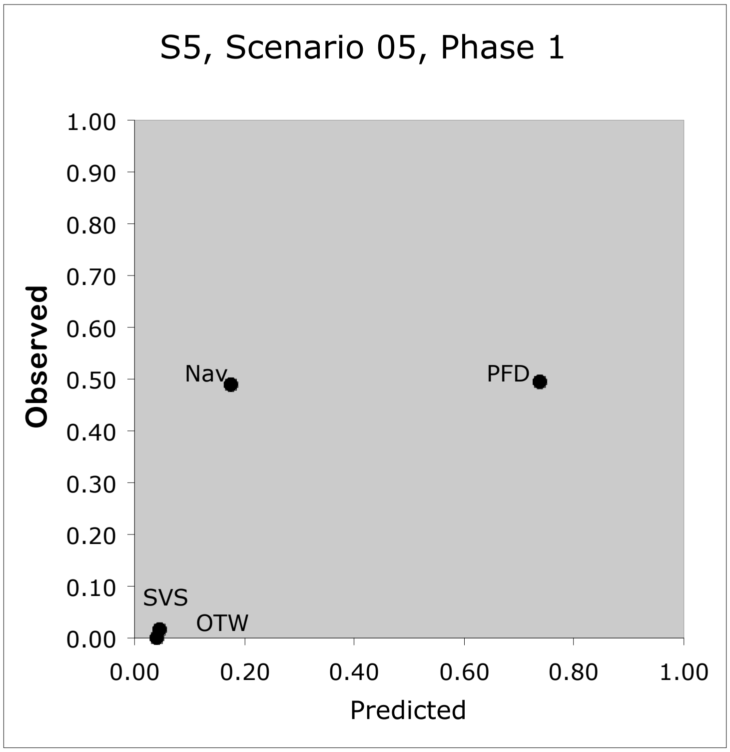


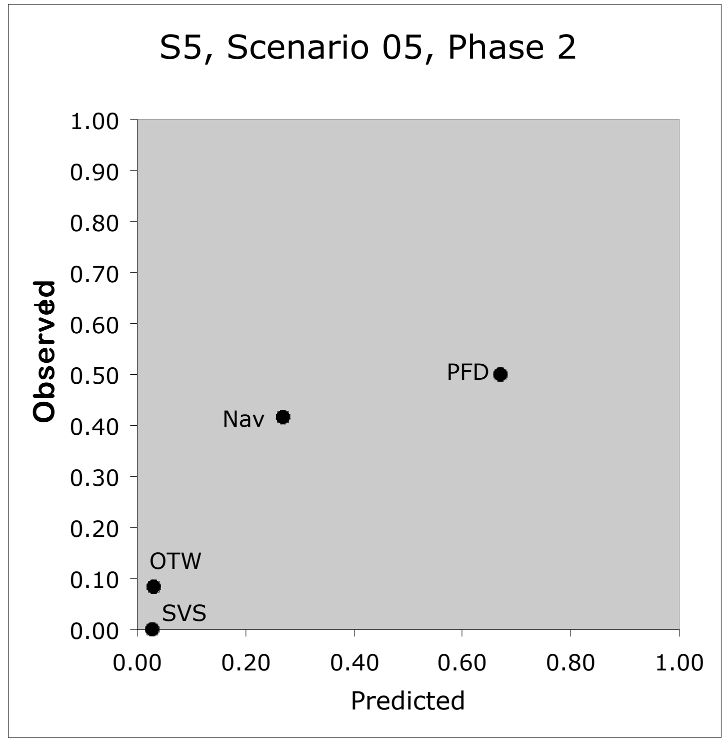


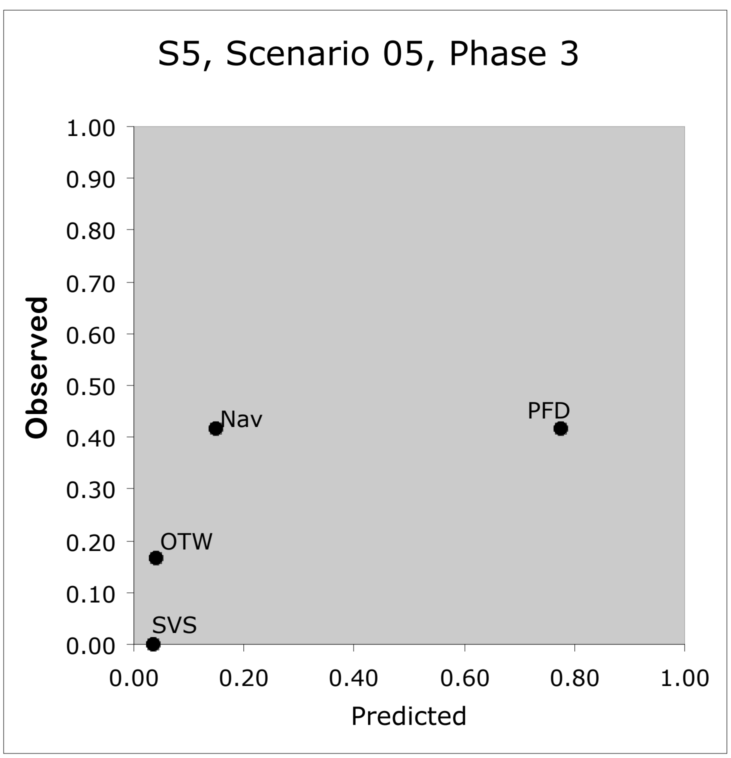


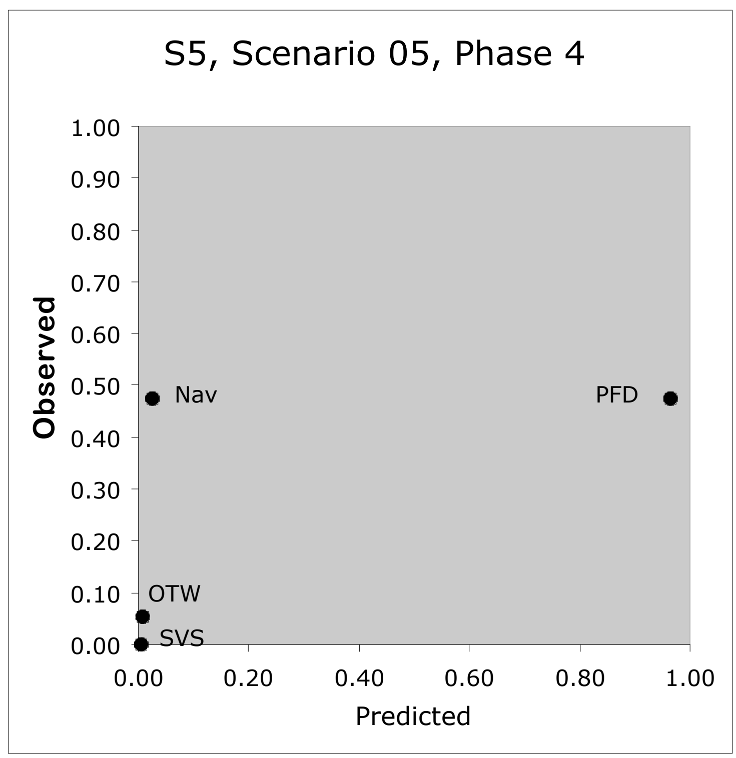


Scenario 5, Subject 5, Phases 1-4

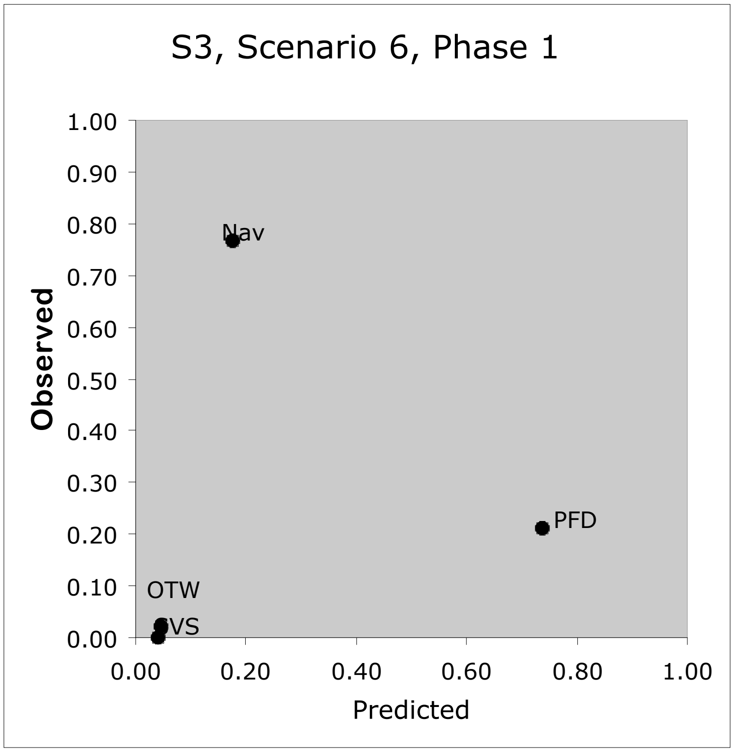


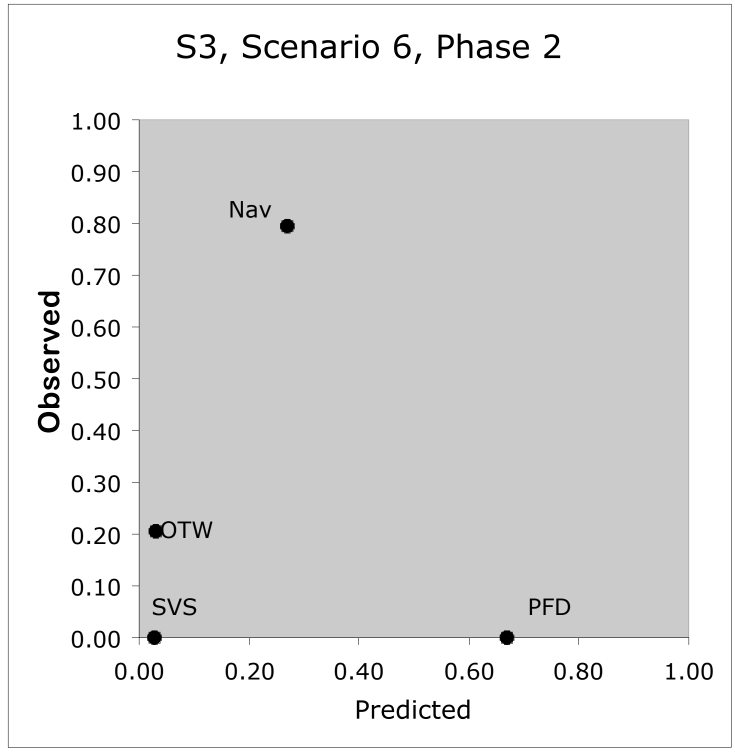


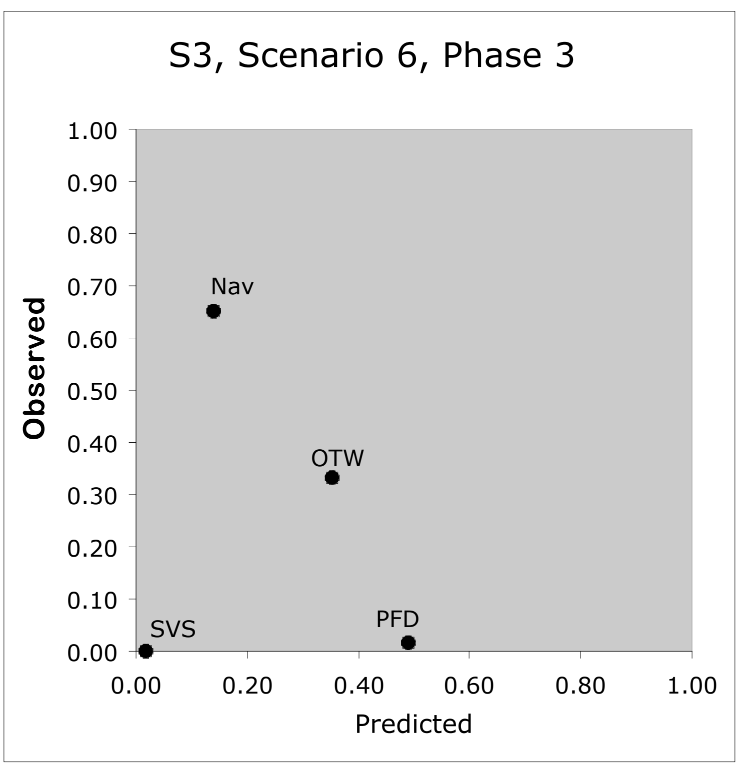


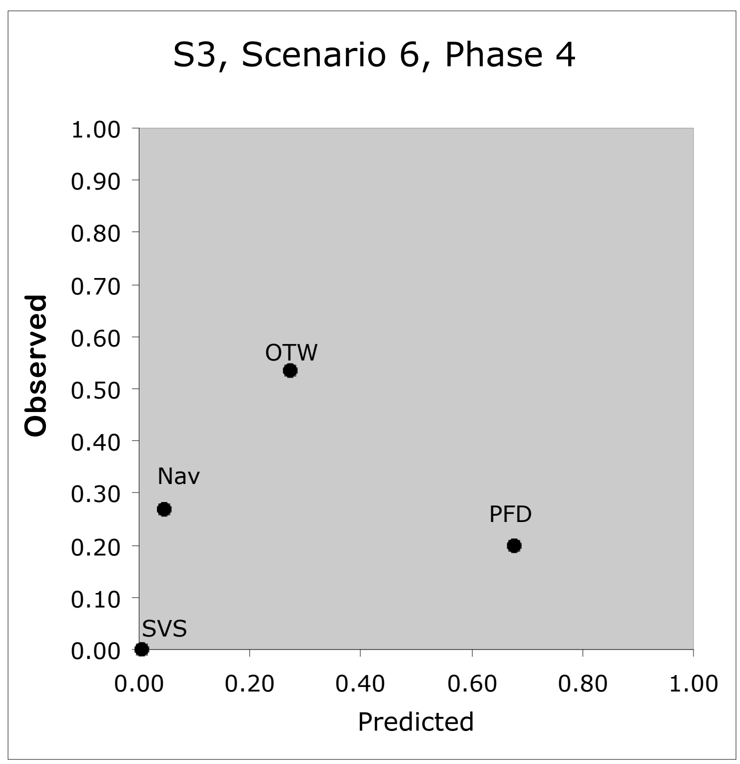


Scenario 6, Subject 3, Phases 1-4

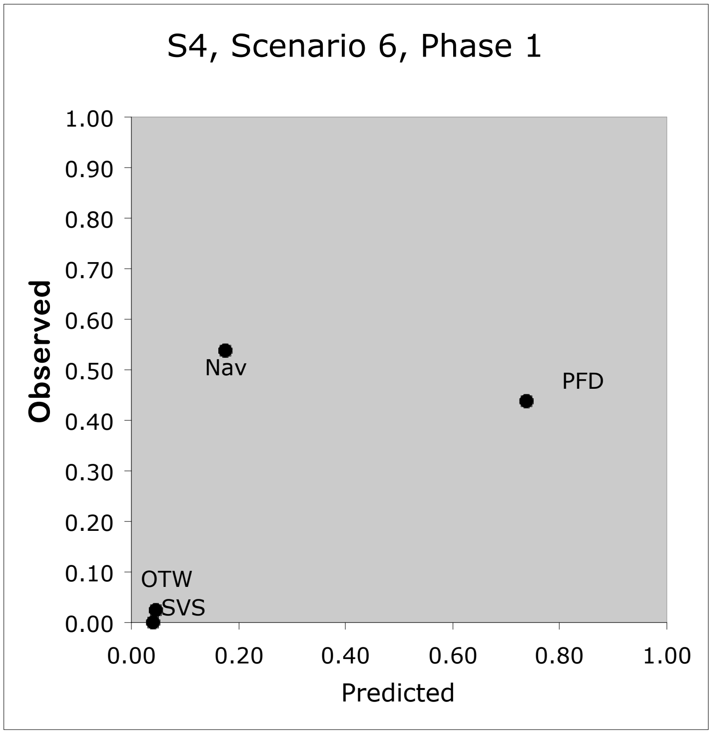


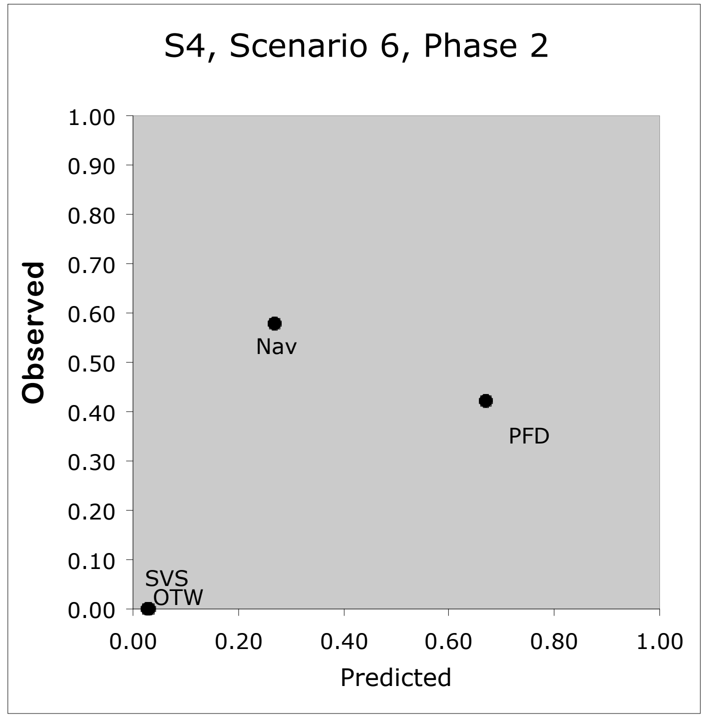


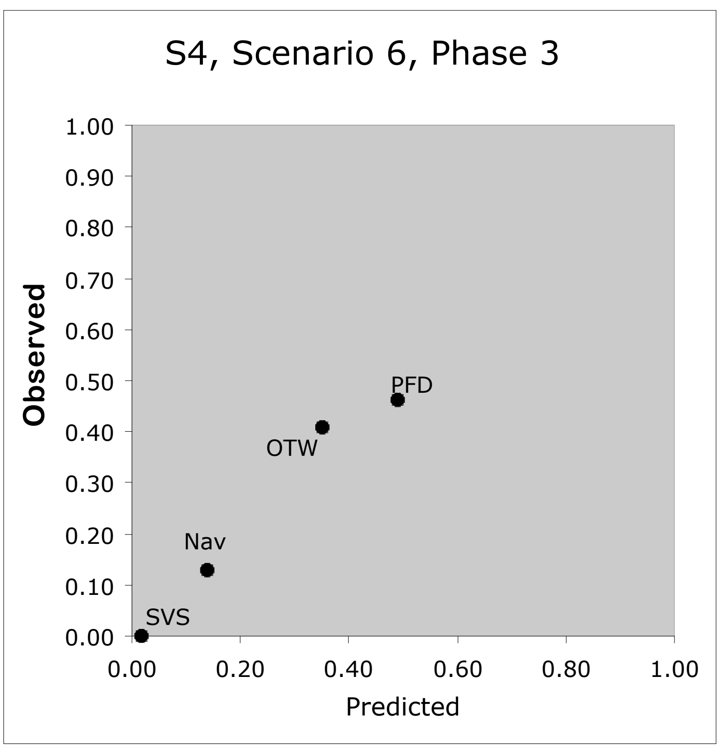




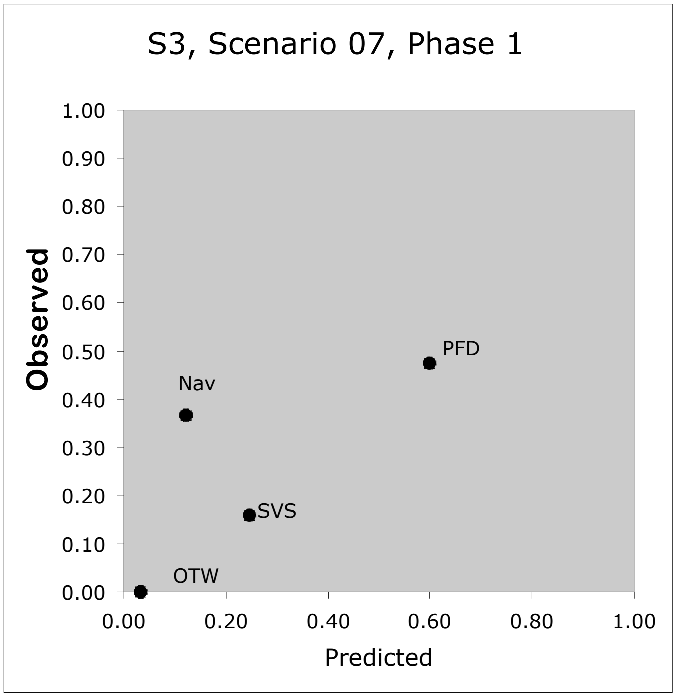
Scenario 6, Subject 4, Phases 1-3

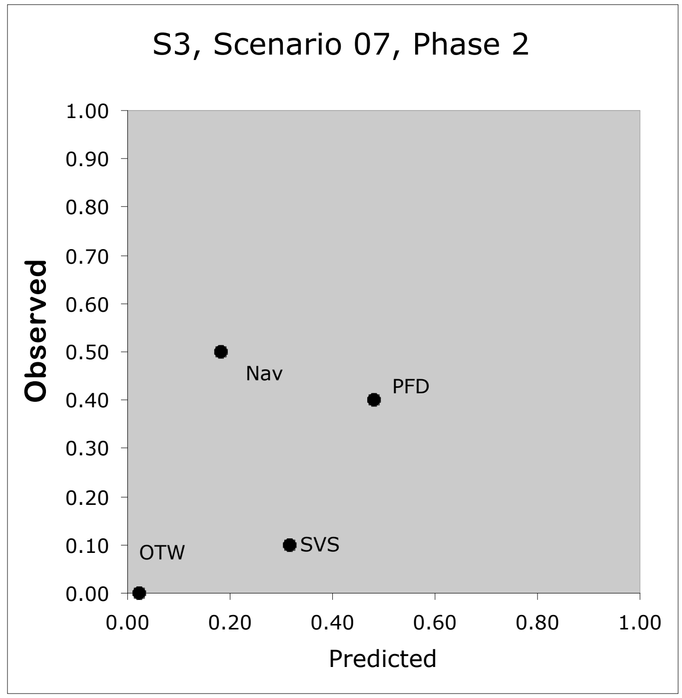


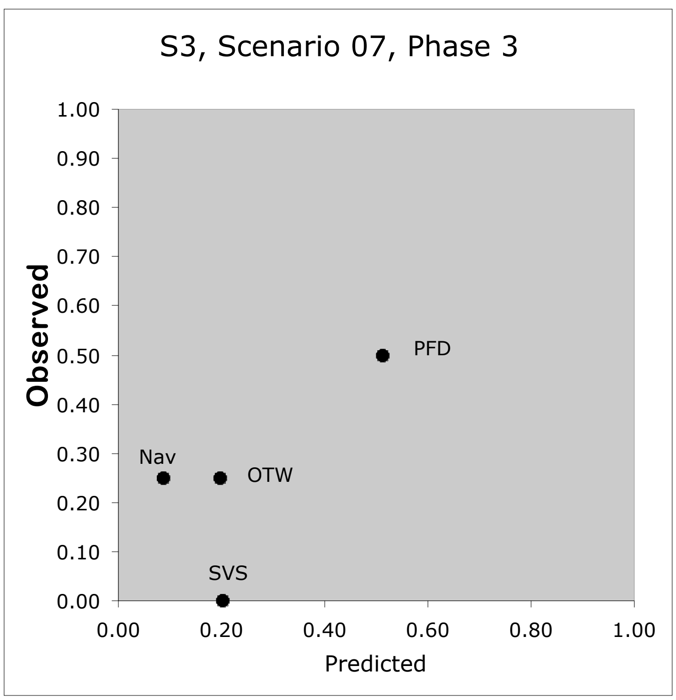


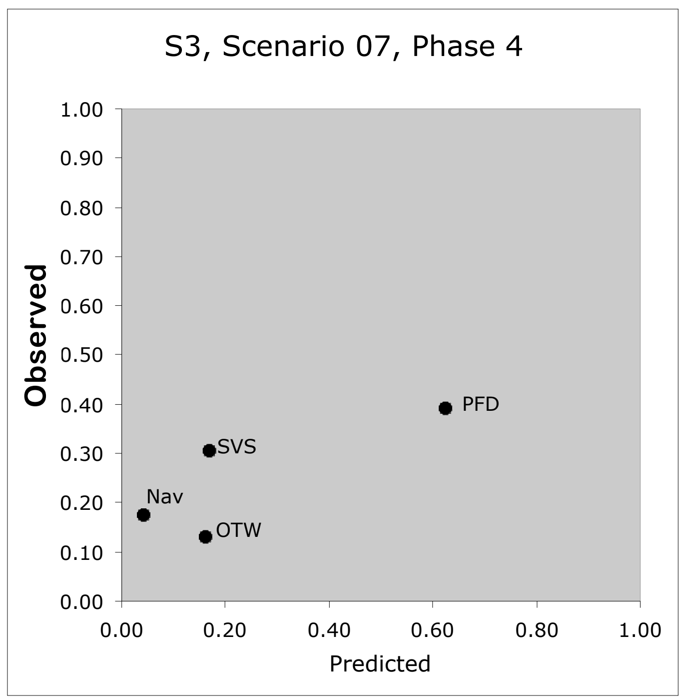


Scenario 7, Subject 3, Phases 1-4

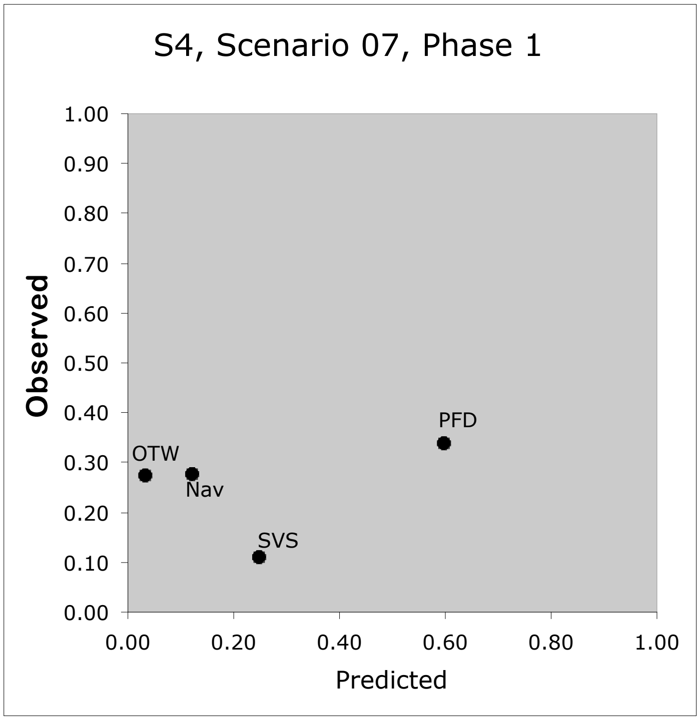


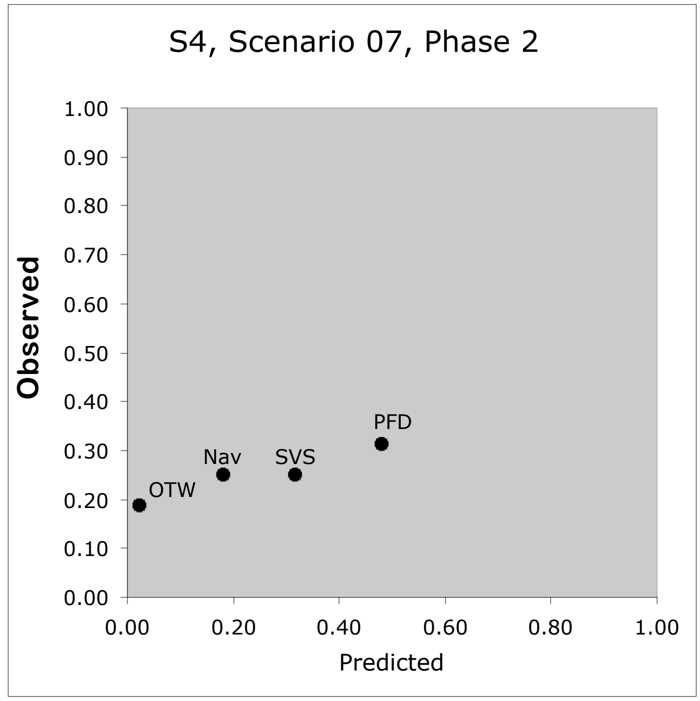


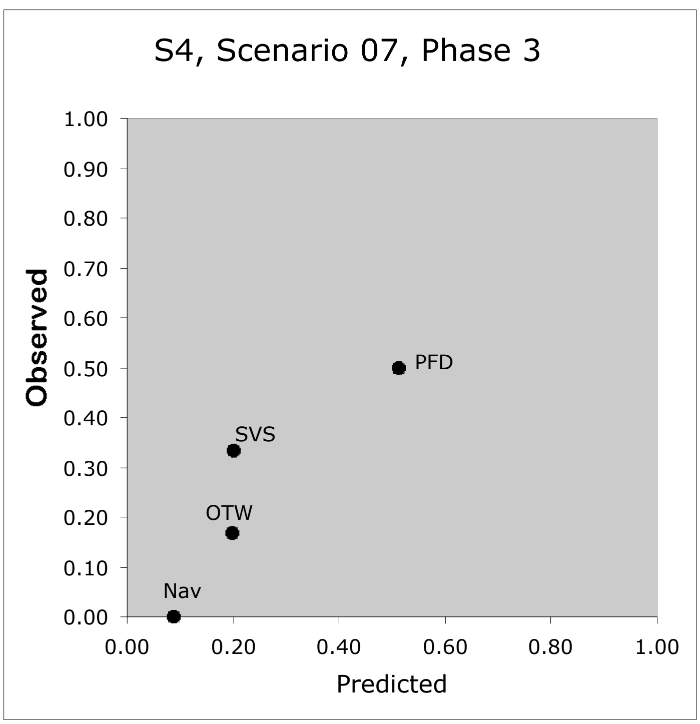


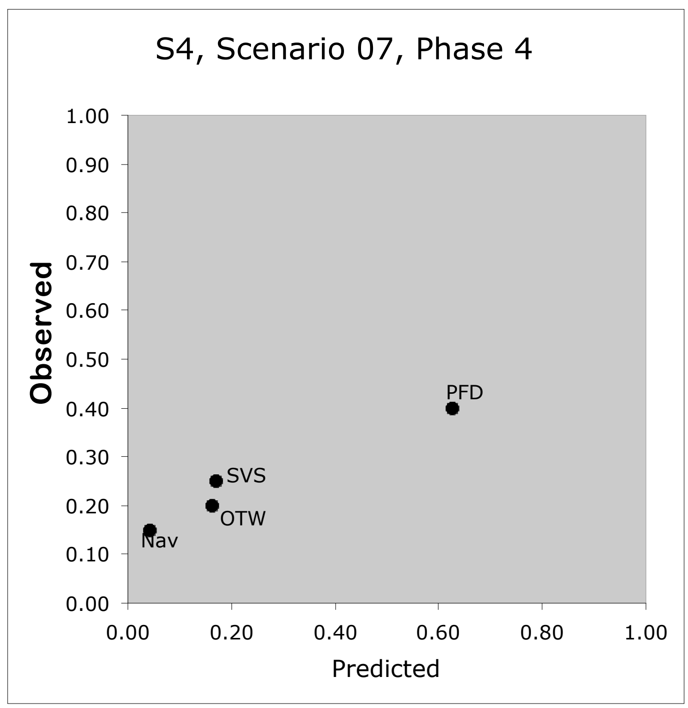


Scenario 7, Subject 4, Phases 1-4

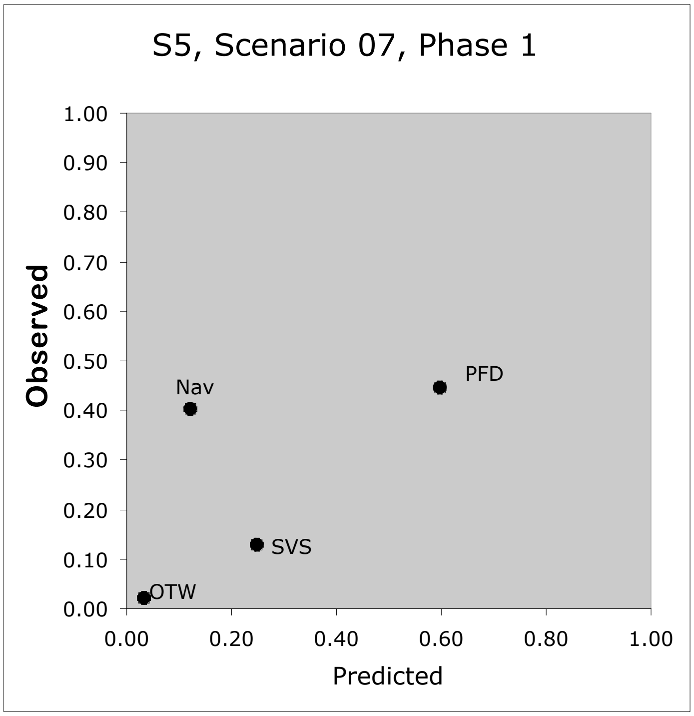


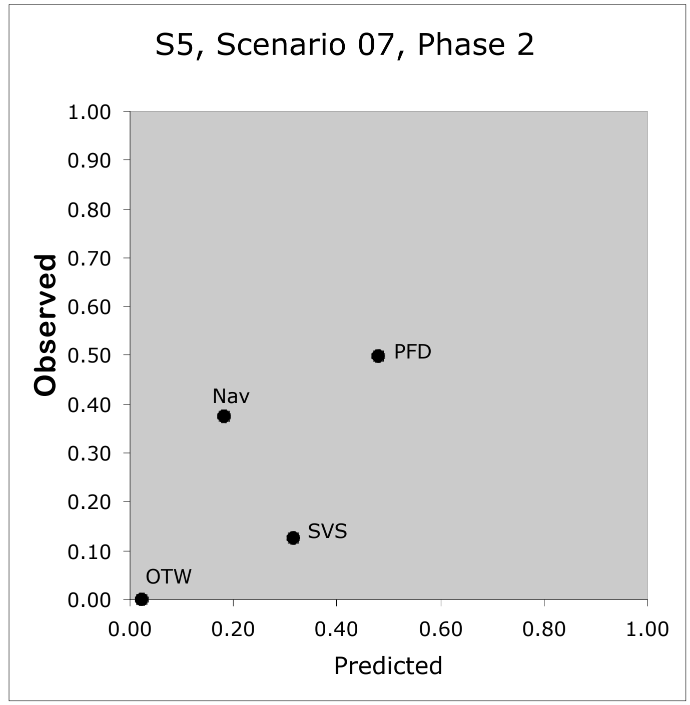


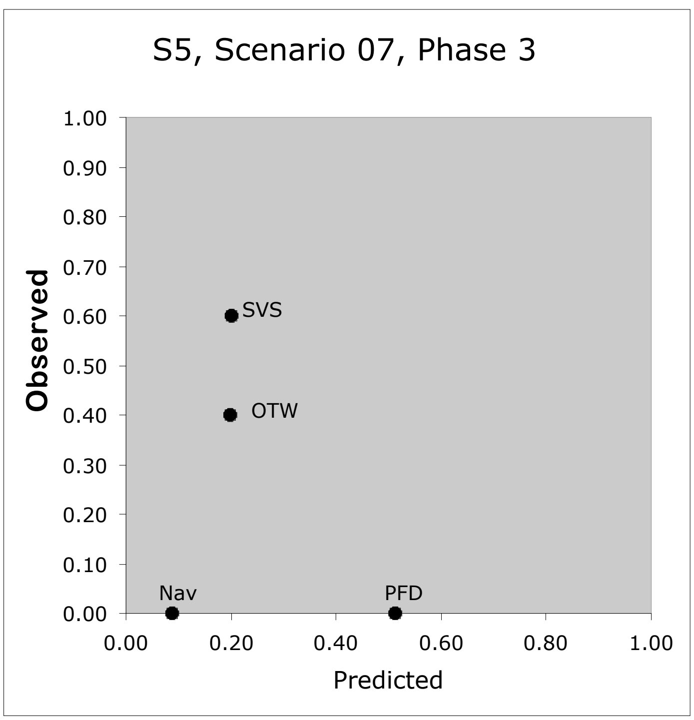


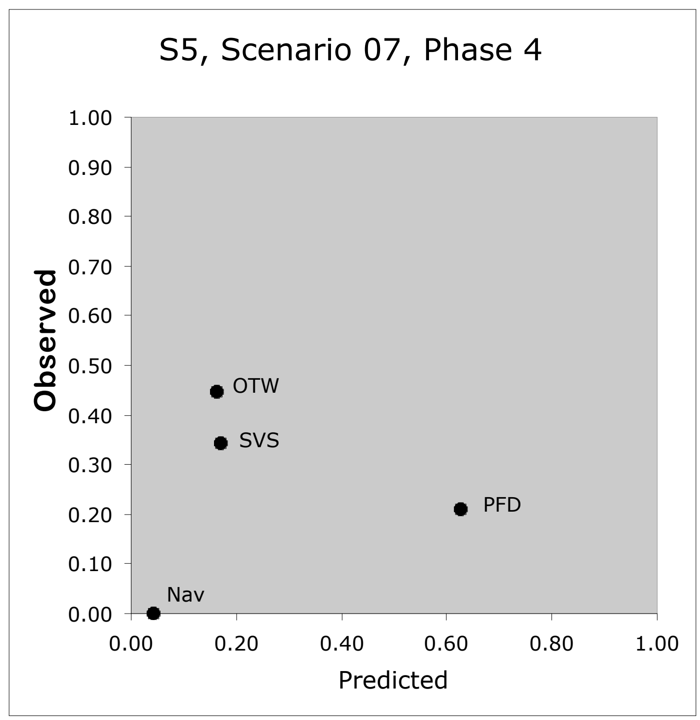


Scenario 7, Subject 5, Phases 1-4

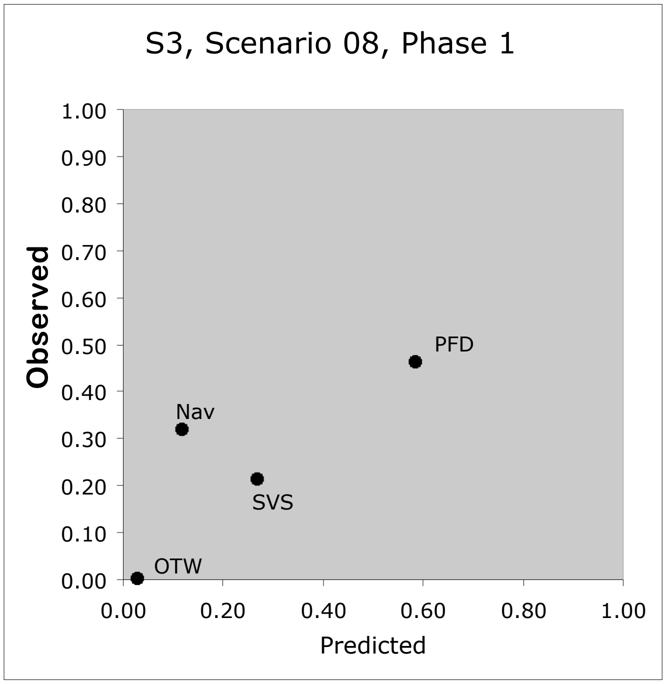


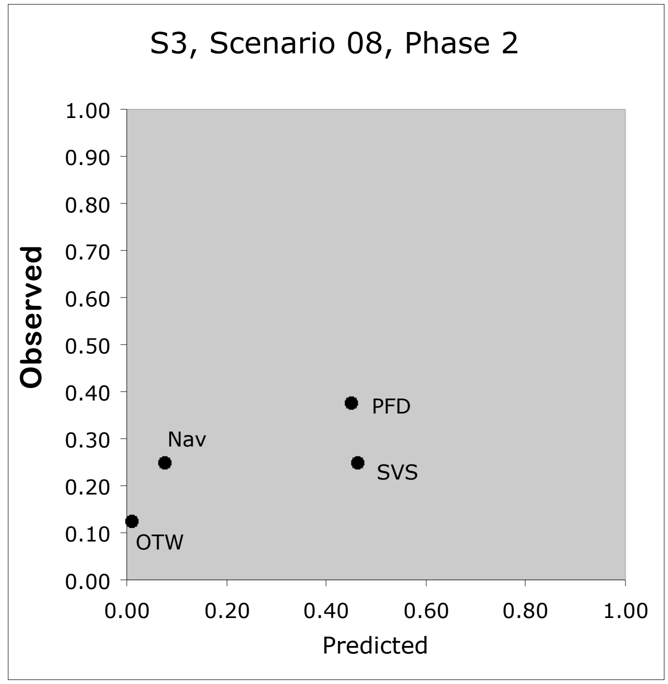


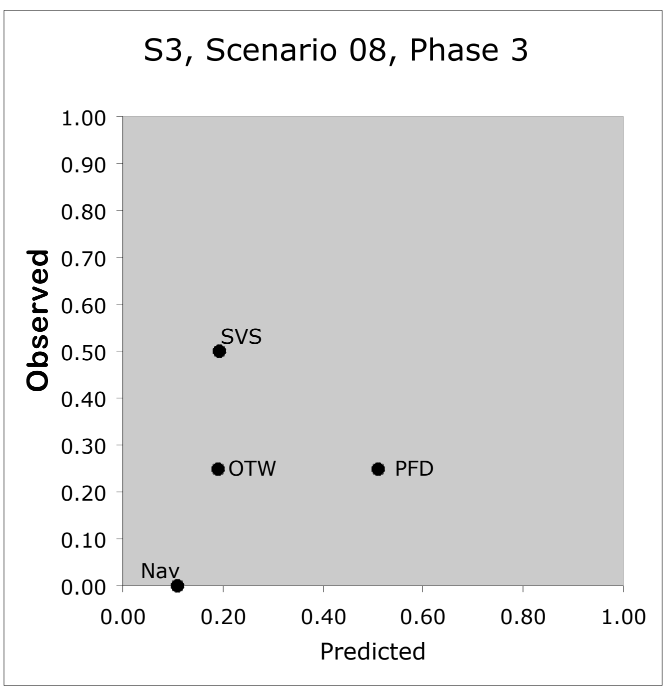


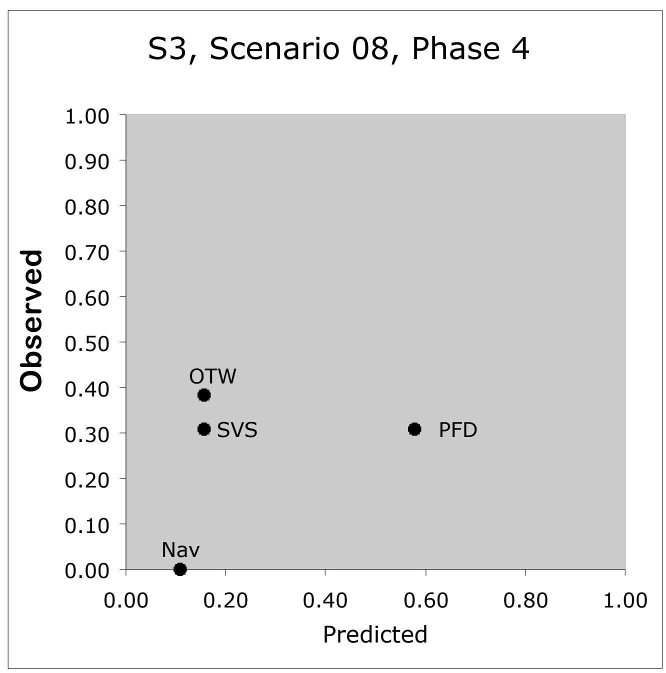


Scenario 8, Subject 3, Phases 1-4

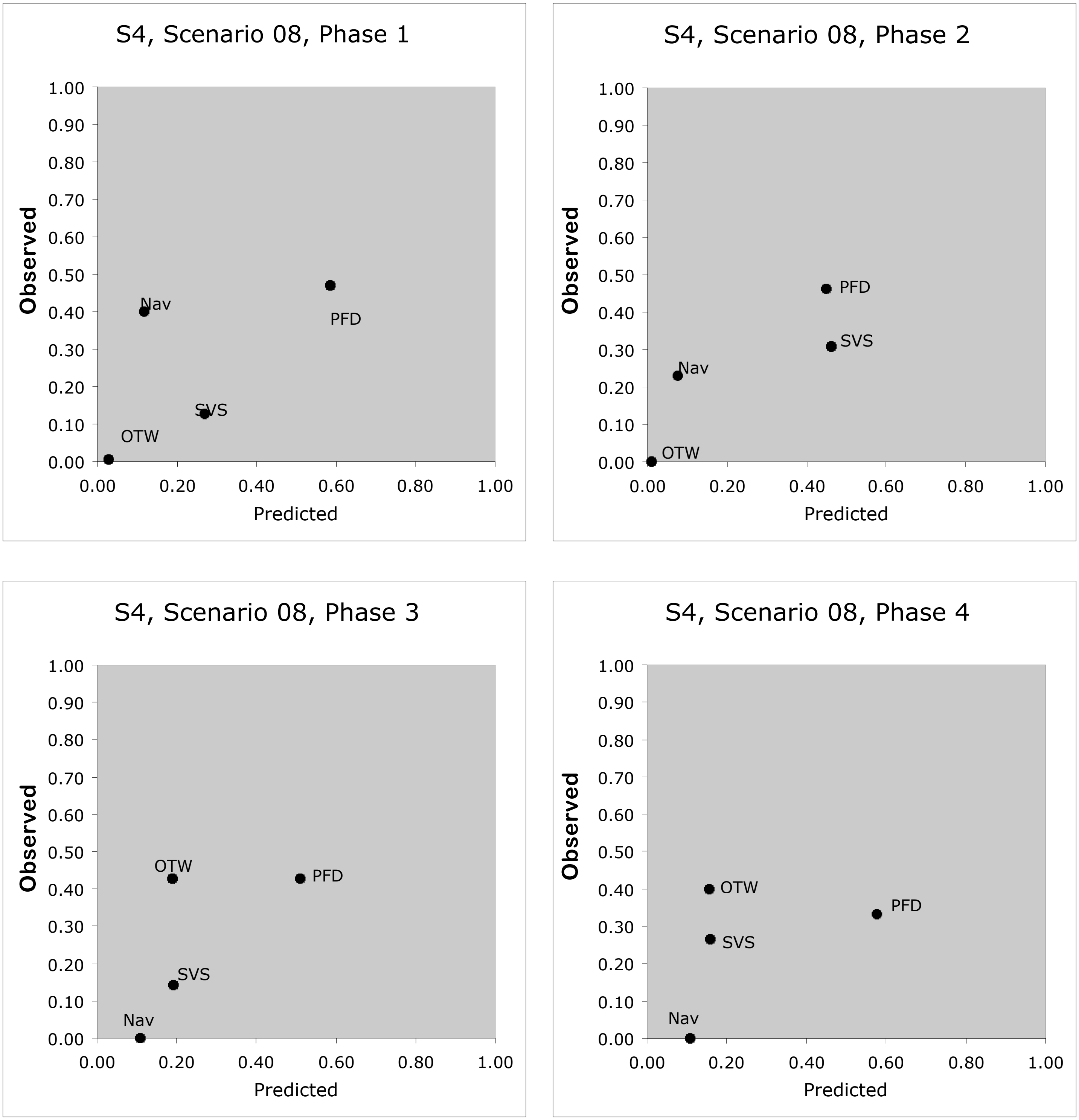




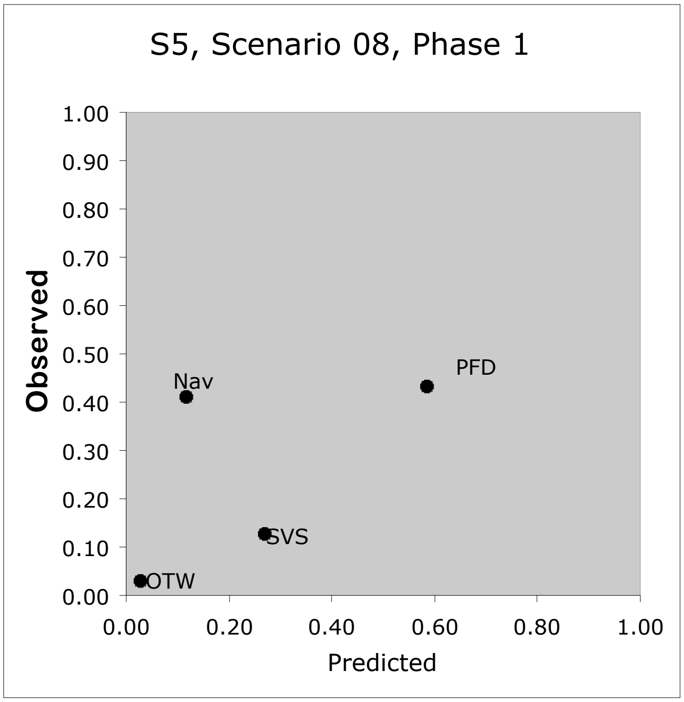


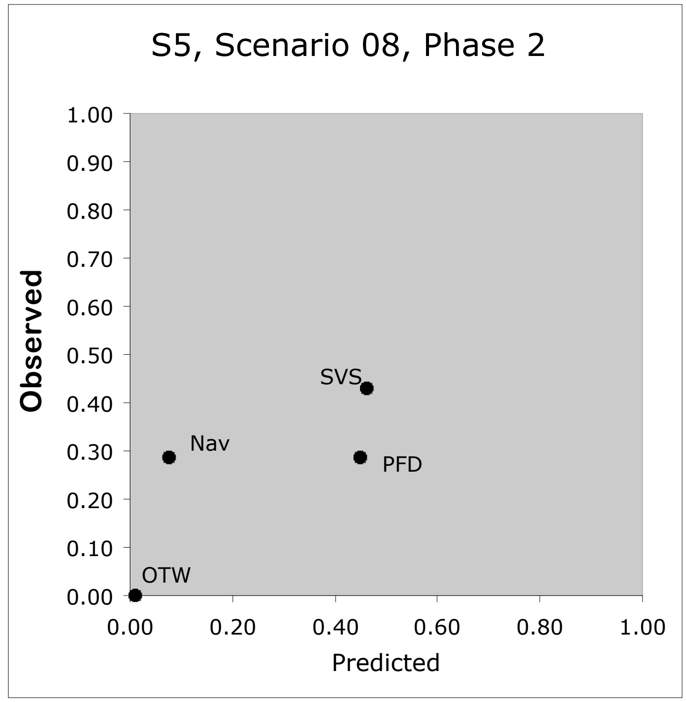


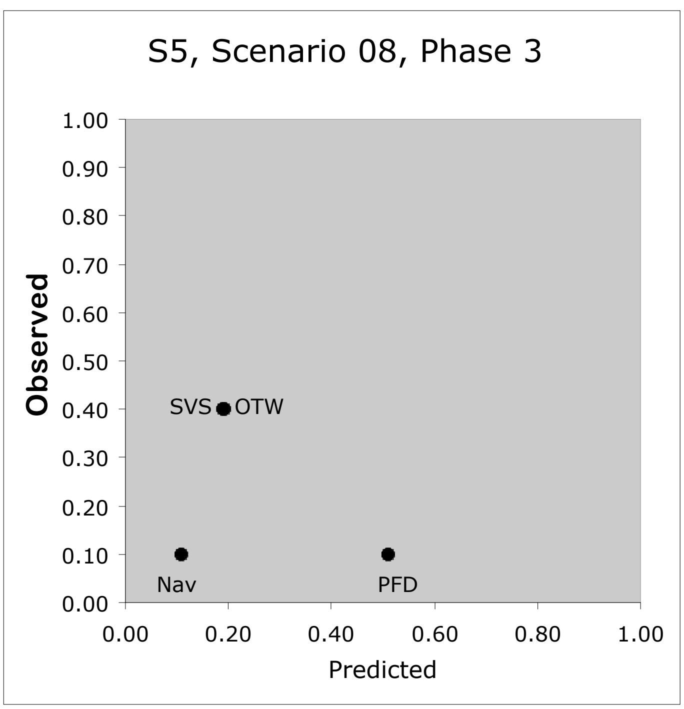
Scenario 8, Subject 4, Phases 1-4

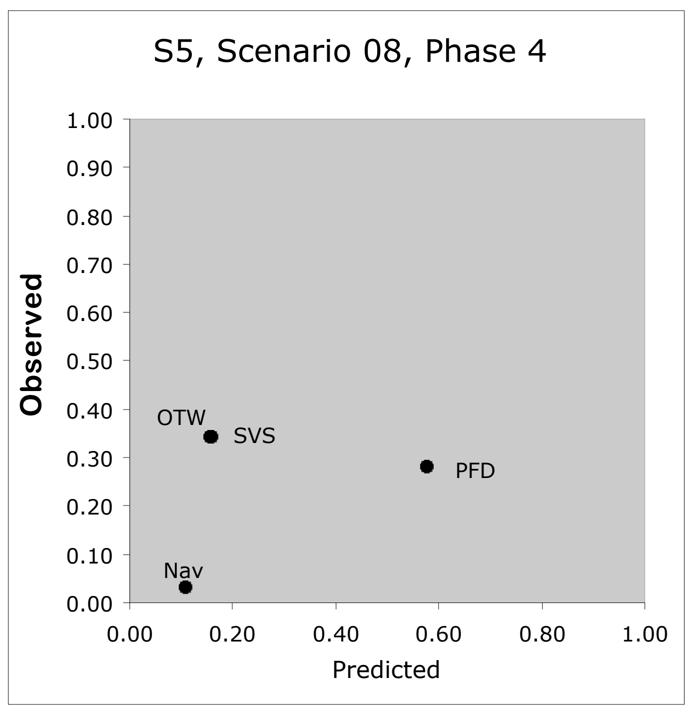


Scenario 8, Subject 5, Phases 1-4

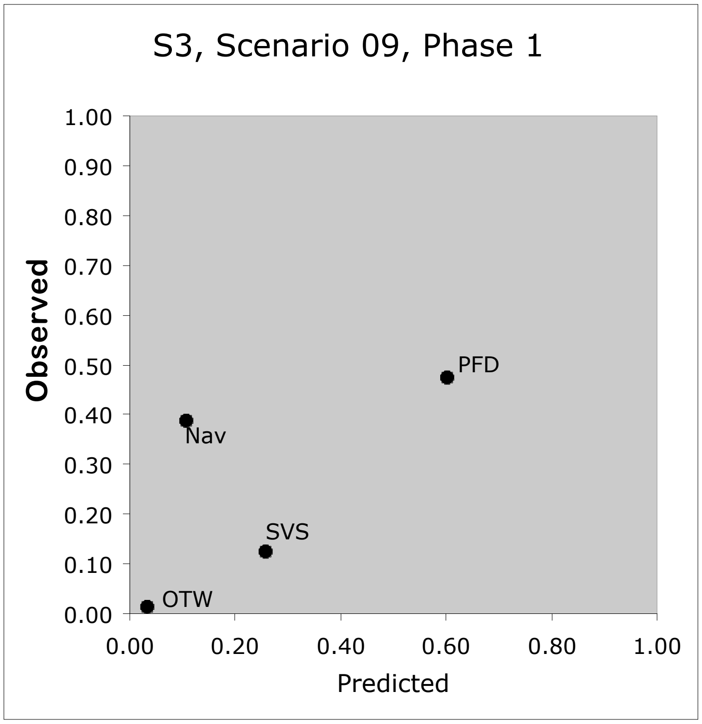


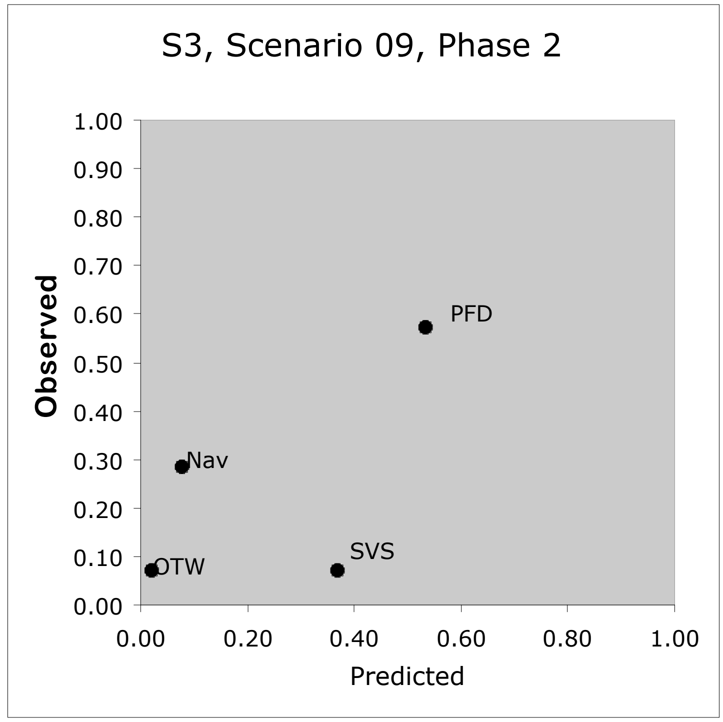


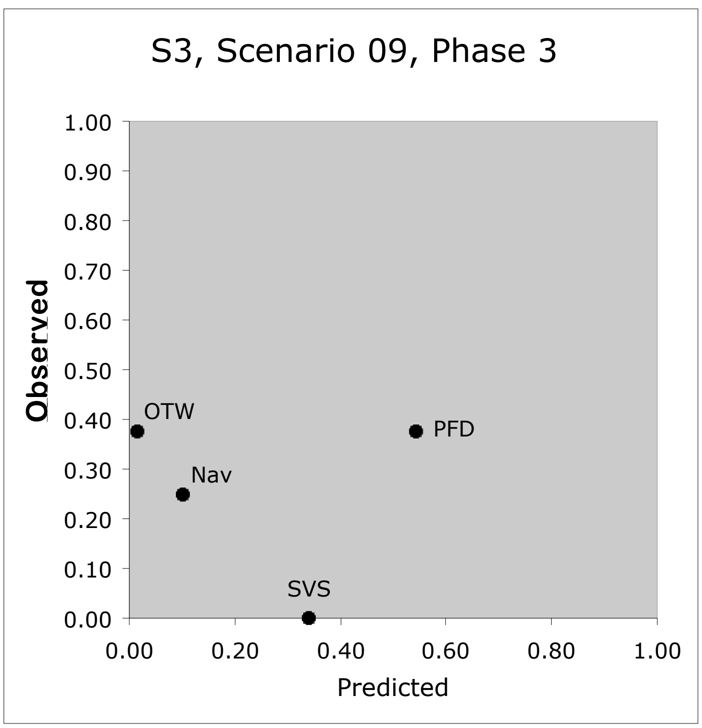


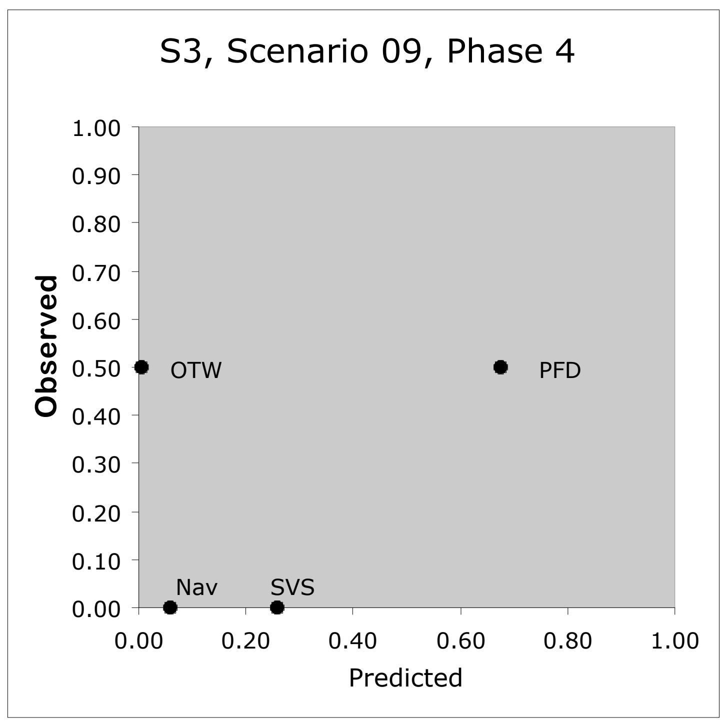


Scenario 9, Subject 3, Phases 1-4

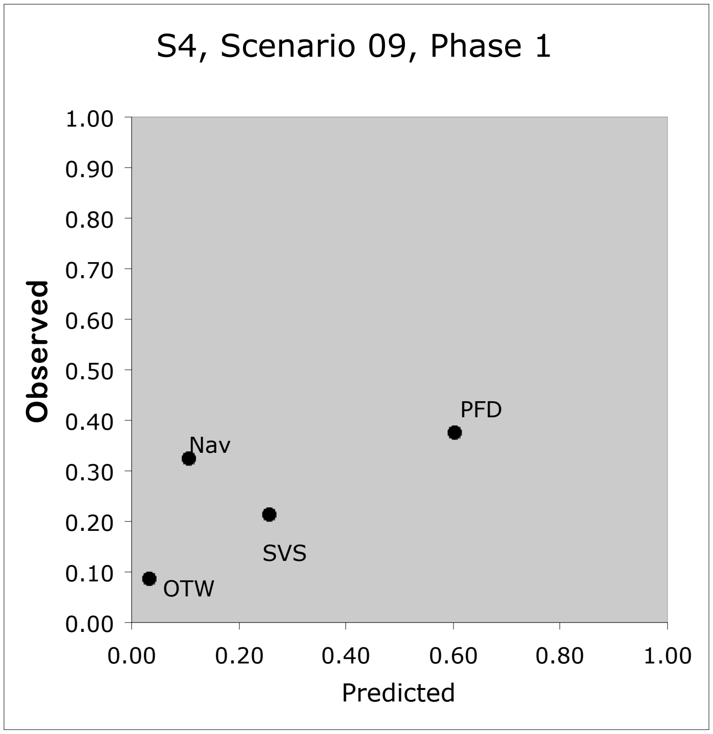


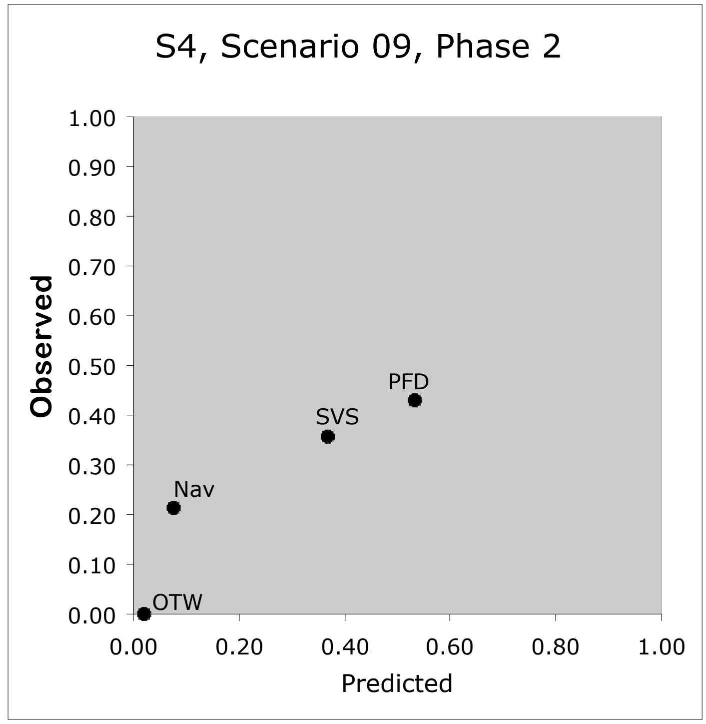


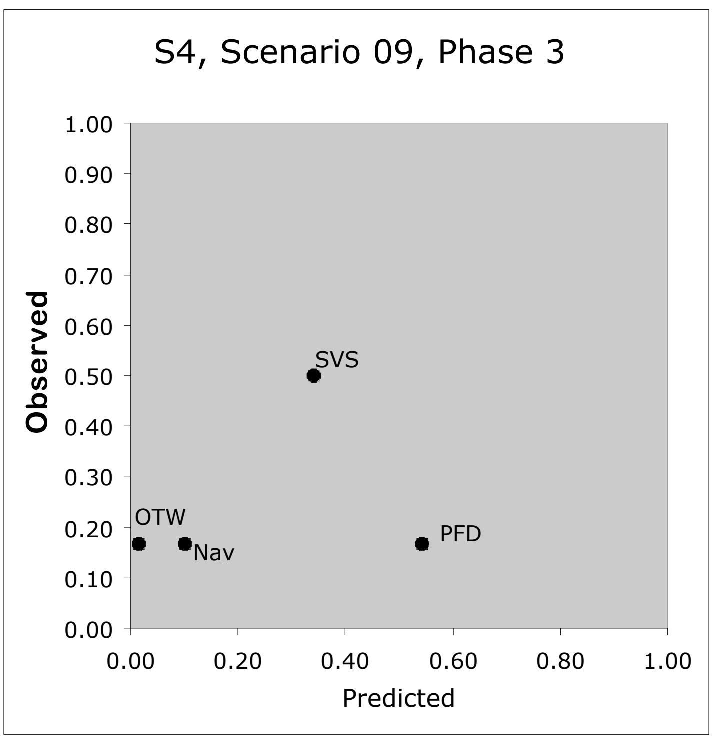


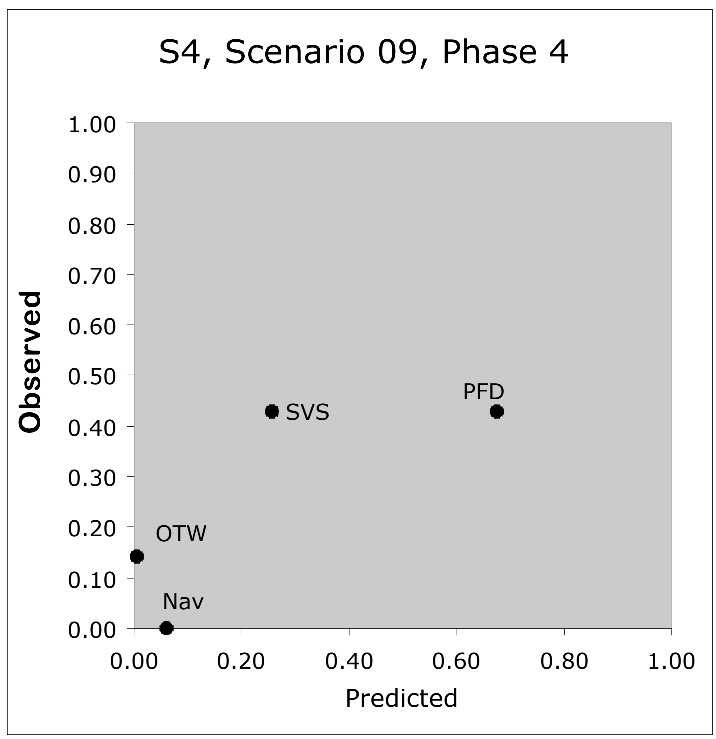


Scenario 9, Subject 4, Phases 1-4

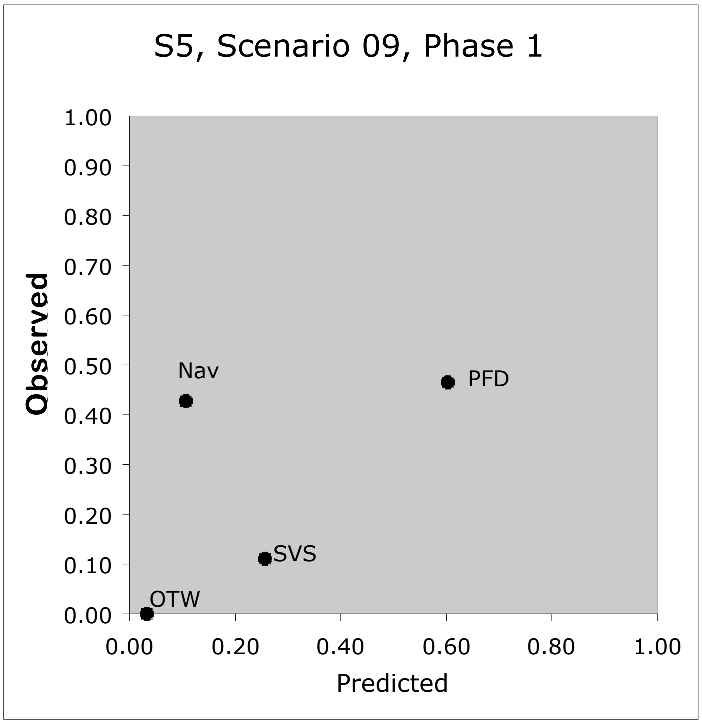


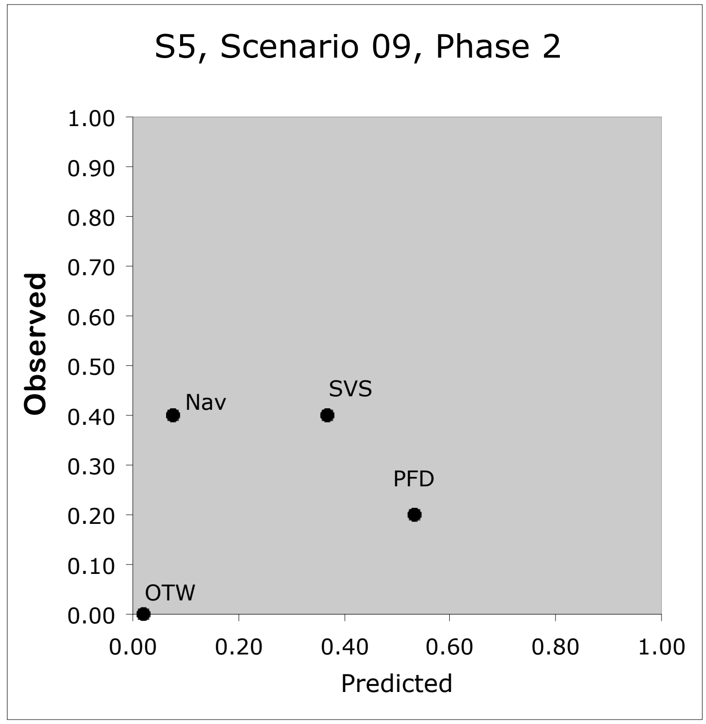


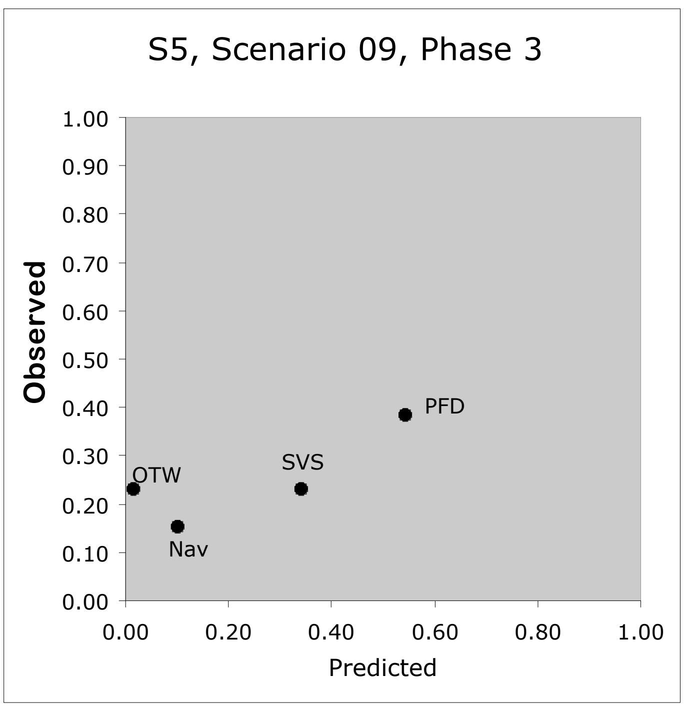


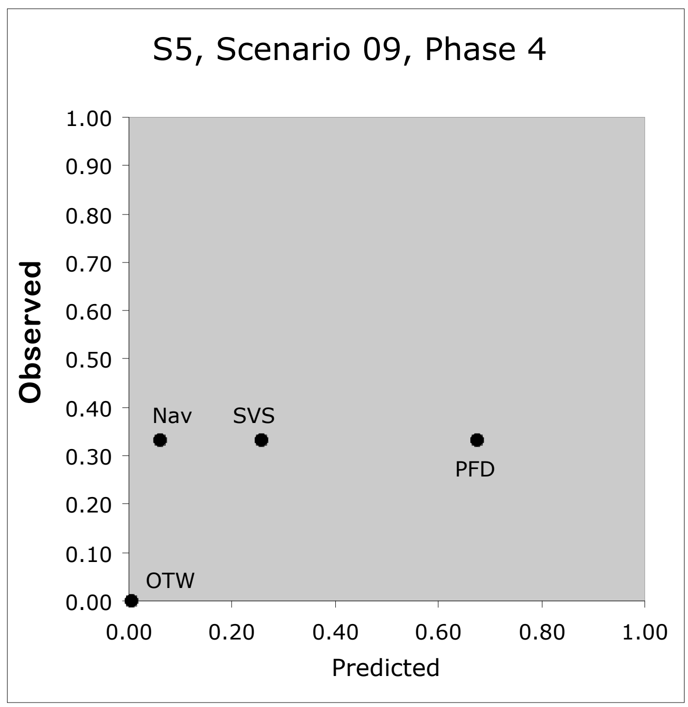


Scenario 9, Subject 5, Phases 1-4

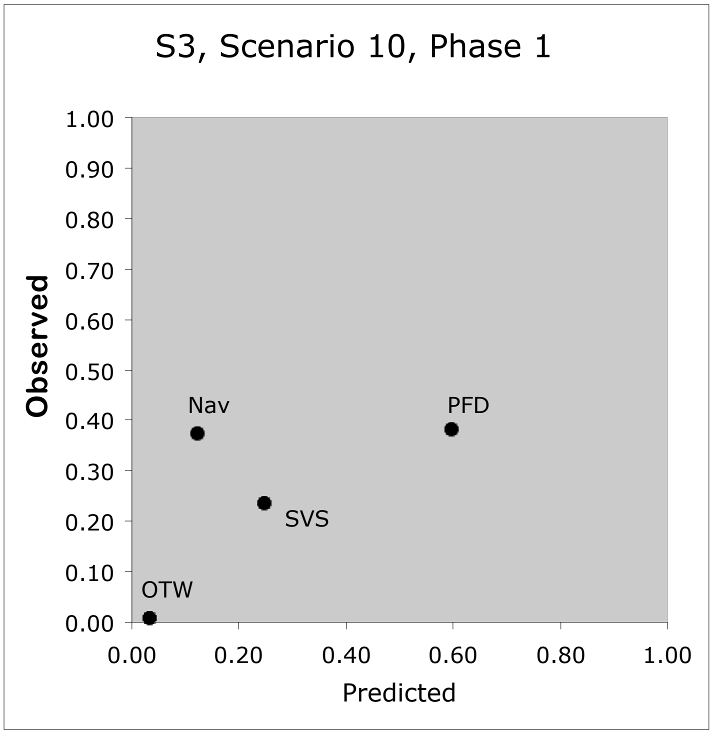


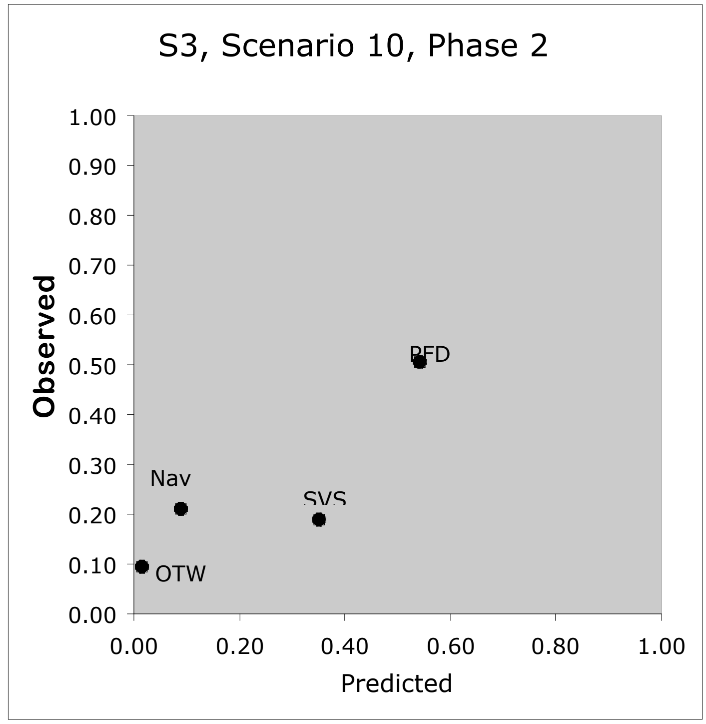


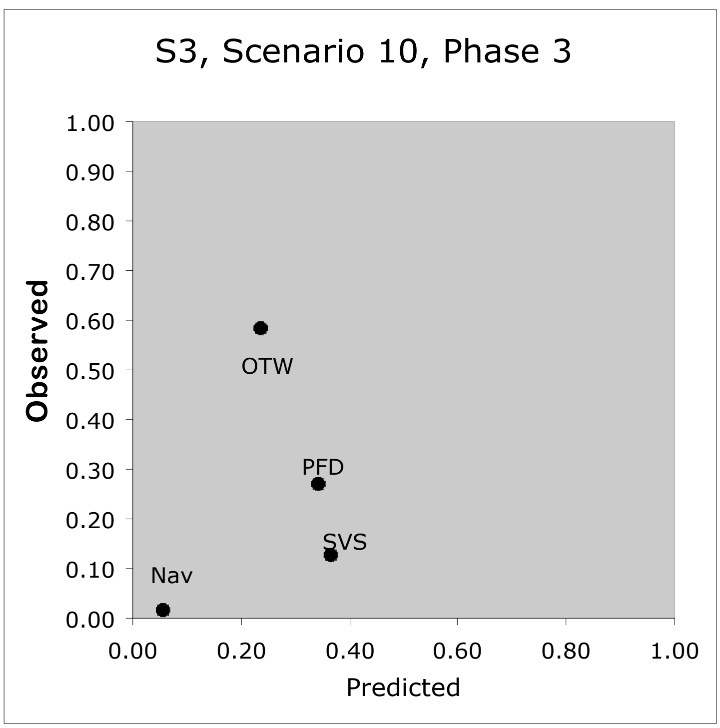


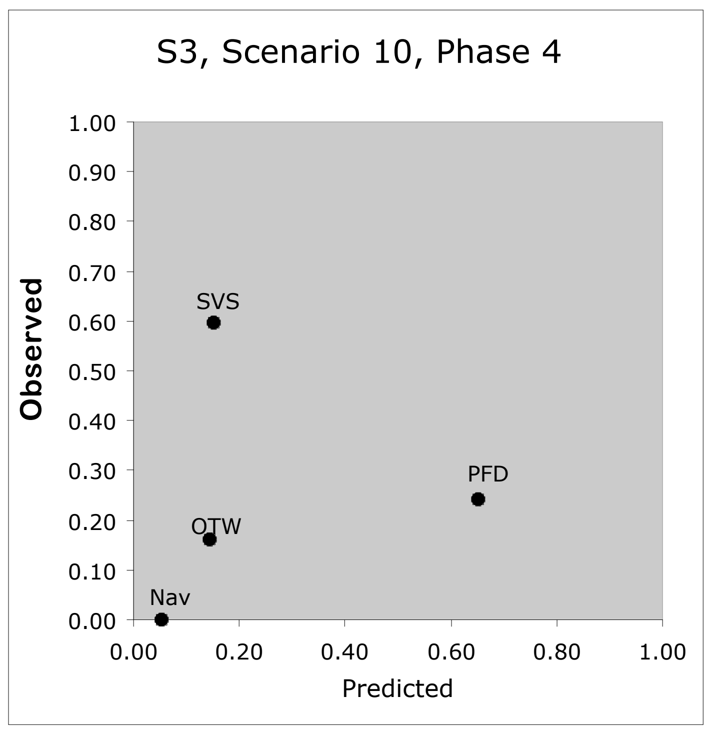


Scenario 10, Subject 3, Phases 1-4

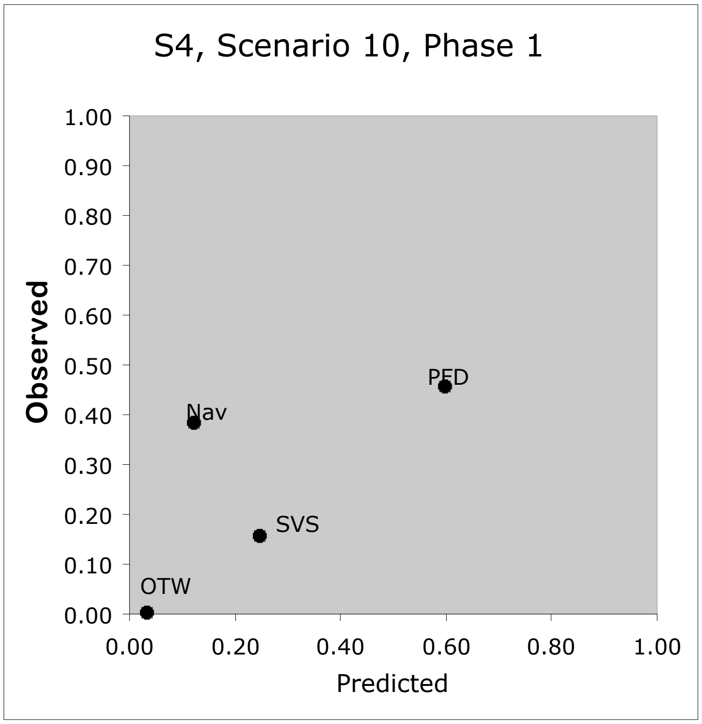


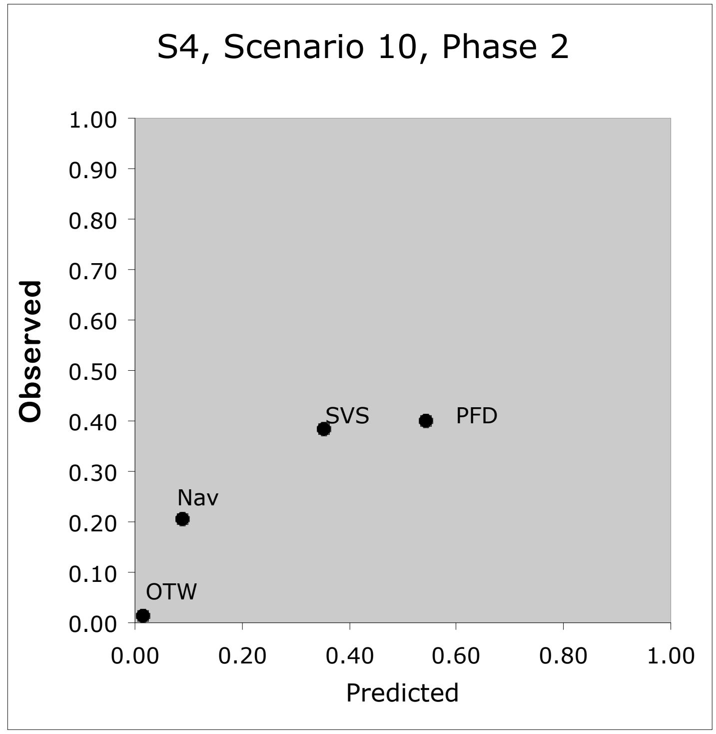


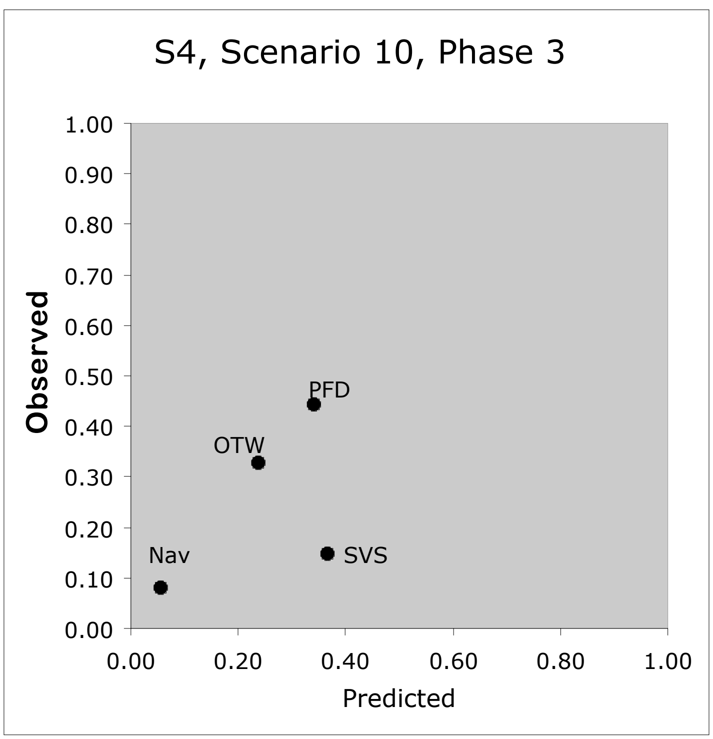


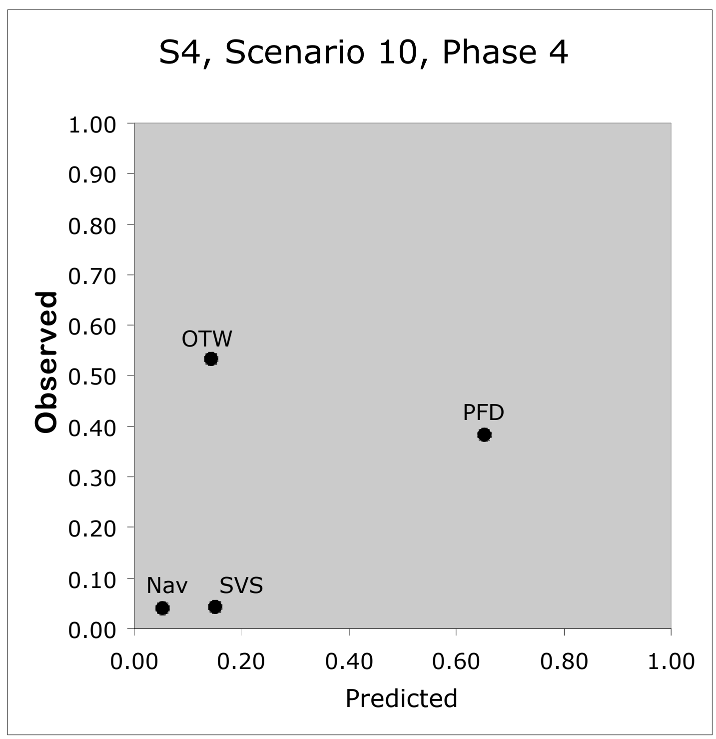


Scenario 10, Subject 4, Phases 1-4

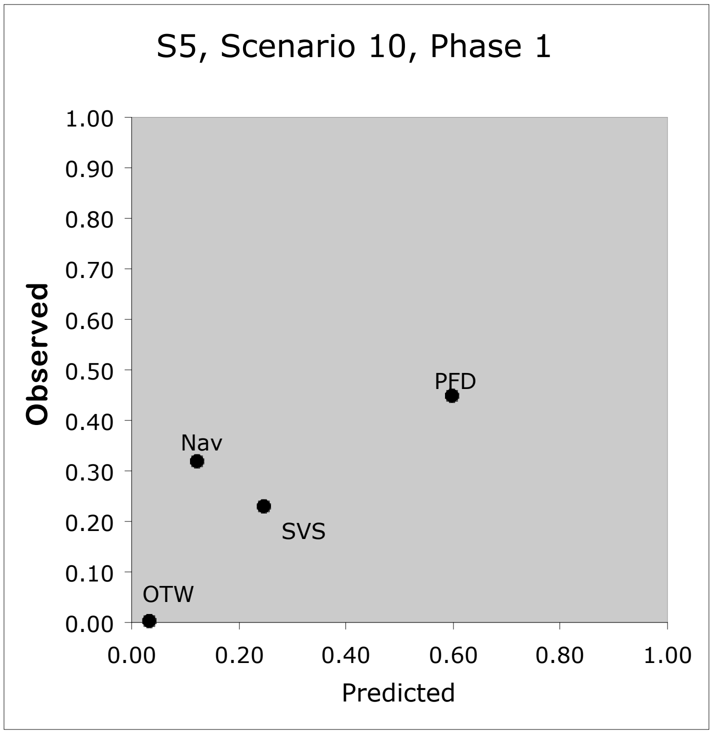


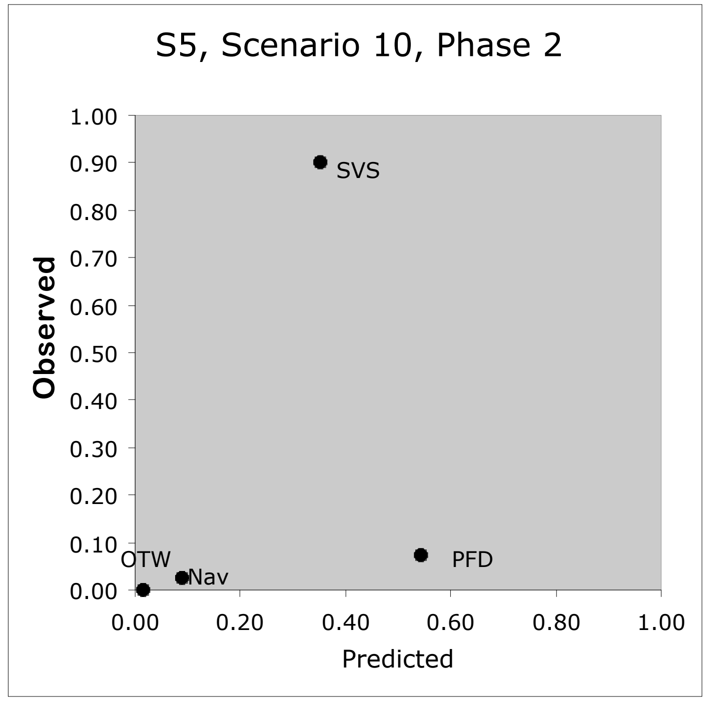


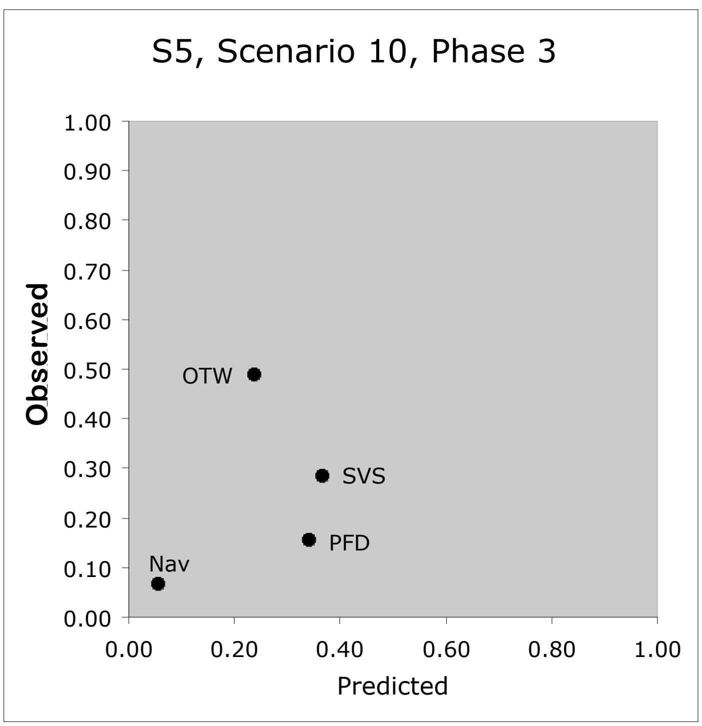


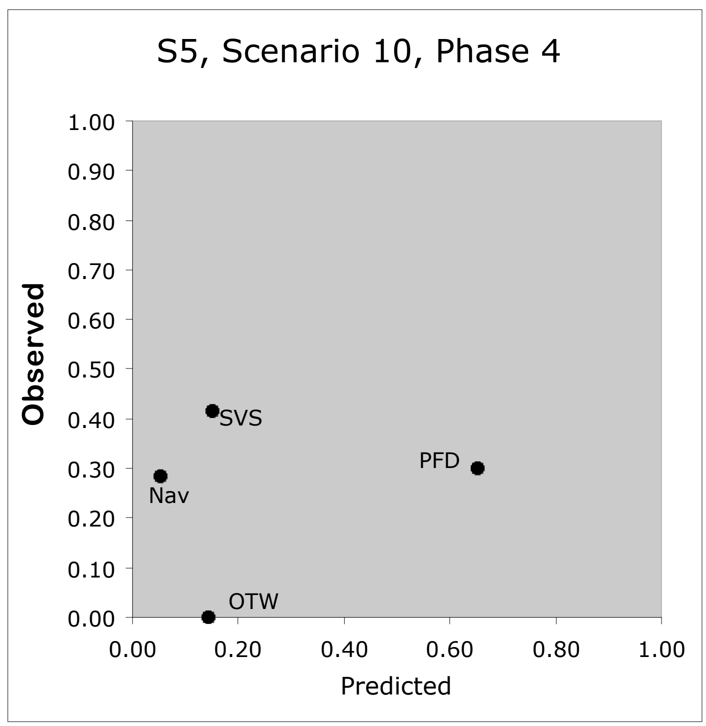


Scenario 10, Subject 5, Phases 1-4









Report Documentation Page

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
September 2003

3. REPORT TYPE AND DATES COVERED
Technical Memorandum

4. TITLE AND SUBTITLE
NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays

5. FUNDING NUMBERS
728-20-10

6. AUTHOR(S)
Edited by David C. Foyle, Allen Goodman, and Becky L. Hooey

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Ames Research Center
Moffett Field, California 94035-1000

8. PERFORMING ORGANIATION REPORT NUMBER
IH-041

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA/CP—2003–212267

11. SUPPLEMENTARY NOTES
Point of Contact: David C. Foyle, M/S 262-4, Moffett Field, California 94035
(650) 604-3053

12A. DISTRIBUTION/AVAILABILITY STATEMENT
Subject Category: 54-04 Distribution: Public
Availability: NASA CASI (301) 621-0390

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The proceedings are a compilation of papers elaborating on presentations made at the 2003 NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays. This conference brought together researchers and modeling teams to present the results of their efforts to use cognitive modeling to predict pilot behaviors during approach and landing operations with a synthetic vision display. An introductory paper reviews program objectives and development strategies underlying a multi-year, multi-thrust effort to advance cognitive modeling into real-world aviation applications. Domain information, qualitative observations, and empirical data which supported model construction are detailed in two papers: a comprehensive cognitive task analysis of approach and landing operations and a part-task study of synthetic vision display usage. Papers describing five separate, but parallel modeling approaches utilizing fast-time simulation for the investigation and analyses of pilot performance in this flight task are presented. Of particular focus in these analyses are predictions of pilot's visual attention. Findings concerning synthetic vision display usage and advances in modeling methodologies are noted.

14. SUBJECT TERMS
Human performance modeling, Cognitive modeling, Visual performance, Synthetic vision

15. NUMBER OF PAGES
228

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500
Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z-39-18
298-102