EXAMINING NEW FLIGHT DECK TECHNOLOGY USING HUMAN PERFORMANCE MODELING

Stephen Deutsch
Richard Pew
BBN Technologies
10 Moulton Street
Cambridge, MA 02138

Aircraft flight deck synthetic vision systems (SVS) always provide a "clear day" view and hence have the potential to improve safety in commercial aviation. Approach, landing, and taxi operations will most readily profit from the SVS capabilities. A part-task simulation study provided data on pilot performance using a baseline and an SVS-equipped flight deck. One effect of adding the separate SVS to the flight deck was that pilot scan patterns changed significantly—more time was devoted to the attitude displays and less to the navigation display. Concern over the change in well-established scan patterns lead to the suggestion that the SVS, an attitude display, be combined with the primary flight display as a single Enhanced-SVS attitude display rather than augment it as a separate display. A human performance model study was used to reproduce the results of the part-task study and then establish that the Enhanced-SVS attitude instrument would restore the original pilot scan pattern.

INTRODUCTION

Commercial aircraft are well equipped to operate in instrument meteorological conditions (IMC), yet there is room to improve safety still further. NASA has designed a Synthetic Vision System (SVS) that provides a small screen "clear day" view under all operating conditions. A part-task experiment was recently conducted at the NASA Ames Research Center in which three commercial pilots served as subjects, each executing a series of simulated approach and landing scenarios using runway 33 left at Santa Barbara Municipal Airport (Goodman, Hooey, Foyle, & Wilson, 2003). Data collected included aircraft state and control inputs, eye-tracker data, and the video recordings from an eye-tracker and a room-view camera. The collected data was provided to several human performance modeling teams that were then asked to use these data to refine their models and extend the findings of the part-task scenario experiment. In this paper, we describe how we made use of our human performance models to address a concern that arose in examining the human subject data and preparing for the human performance model trials.

The NASA part-task experiments included ten scenarios that explored three independent variables: (1) a *display* condition that included a baseline flight deck configuration and a flight deck supplemented with a captain's SVS; (2) a *visibility* condition that included visual meteorological conditions (VMC) and instrument meteorological conditions (IMC); and (3) a *situation* condition of four distinct approach scenario trials. The approach situations included a nominal landing, a landing that included a late air traffic controller request to switch to an adjacent parallel runway, an attempted landing that required a missed approach, and an approach during which the SVS was misaligned. In the SVS-misaligned scenario, the view provided by the SVS was readily determined to be offset from the out-the-window view as the aircraft emerged from the cloud cover (Goodman et al., 2003).

The simulation baseline flight deck included a side-stick controller, a mode control panel (MCP), a primary flight display (PFD), and a horizontal situation indicator (HSI). The side-stick controller and the MCP provided the subject with control of the flight path of the aircraft. The PFD provided attitude information (heading, speed, altitude, and altitude rate) and annunciators for the MCP mode. The HSI provided navigation information showing current location and flight path waypoints. When the SVS was present, it provided attitude information similar to the PFD in addition to the "clear day" view of the terrain.

A review of the eye tracking data yielded several interesting findings that provided the basis for the approach taken in the subsequent human performance modeling effort. Clearly three subjects are a modest sample, yet the nature of the adaptation to the SVS scenarios proved very interesting. There were broad individual differences in the use of the SVS. One subject made modest use of the SVS when it was available; one subject made extensive use of the SVS; and the third subject fell approximately midway between the other two. In part, this can be explained by the considerable overlap in functionality between the SVS and PFD. Both provided very similar basic attitude information. In the case of the SVS the attitude ball was replaced by the "clear day" terrain view, but it did not include the flight mode annunciators.

The second finding, in this case consistent across the three subjects, was a shift in the balance between the use of the attitude display(s) and the navigation display (the HSI) when the SVS was included in the configuration. When the SVS was present, time devoted to the attitude displays (the PFD and the SVS) increased at the expense of time devoted to the navigation display (the HSI). The shift in the allocation of time between the attitude and navigation displays was investigated in the human performance modeling study.

The decision to employ human performance models in this study had a central role in determining the course of the investigation. It was at the point of preparing to add the SVS to the flight deck model that its impact on the allocation of scan time across the instruments first surfaced. In thinking about how to model aircrew procedures for this configuration, it became clear that the addition of the second attitude display would impact the allocation of time among the flight deck instruments; the question was: Just what would that impact be? Hence, we re-examined the part-task experiment data and the second finding related above was identified. One potential path to re-establishing the standard PFD scan pattern was to provide an SVS that subsumed the functions of the PFD, so that it could be used as the single flight deck attitude instrument. The modeling environment enabled us to readily explore this possibility. A model for the new Enhanced-SVS was developed and pilot procedures were adapted to use the newly modified flight deck.

The D-OMAR Model

To support the modeling aspect of the project we used the BBN-developed Distributed Operator Model Architecture (D-OMAR), a general-purpose discrete-event simulator. D-OMAR 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—provide the basis for constructing the alternate model architectures. The particular models employed for this NASA research task were a further development within an architecture for human performance modeling that has been evolving 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. Rather than being a particular cognitive architecture, D-OMAR was developed 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 through several projects to address aircrew research problems. We view this level of flexibility in model architecture as essential to the effort to improve the capabilities of and expand the uses of human performance models.

One of the principle areas of research in the development of D-OMAR has been in the modeling of 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 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 use of the pilot's 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 aircrew 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 capabilities-based 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 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 a task. The publish-subscribe protocol also serves to move information among the functional centers.

From a top down perspective, the things that a person knows how to do, basic person skills (e.g., coordinated hand-eye actions to set a flight deck selector) and domain specific skills (e.g., the captain 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. These top-level 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-procedure hierarchy and linked by the publish-subscribe protocol. A subset of the procedures is active; most are in a wait-state. The procedures in a wait-state represent the capabilities to complete actions currently underway and to respond to impinging events.

Within this framework, process (Edelman, 1987; 1989) can be seen to have 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.

Several analysis tools are available in the D-OMAR simulation environment to assist in the analysis of the model results. An online trace was structured to provide a high level view of the execution of the significant steps in the operating procedures. The review of the trace confirmed the execution of the procedures as defined by the cognitive task analysis (Keller & Leiden, 2002a; 2002b). A Gantt-style display of procedure execution for each of the human performance models enabled the detailed examination of the components of the behaviors of the models down to level of task interruptions in the model's multi-task behaviors. Using these tools, it was possible to verify that the behaviors produced by the models were representative of those produced by the commercial pilots in the part-task experiment.

METHOD

We have developed and refined the D-OMAR human performance modeling environment over a number of years and were able to make use of an existing approach and landing scenario that supported a previous study (Deutsch & Pew, 2002). We used a cognitive task analysis provided by Keller and Leiden (2002a) to guide the refinement of the nominal approach and landing procedures and then extended the procedures to include the use of the SVS (Keller & Leiden, 2002b). The cognitive task analyses were developed to describe a real operational environment rather than that employed for the part-task simulation scenarios. In like manner, the procedures that were developed for the captain and first officer human performance models were real-world operational procedures. Flight deck procedures and instruments were functionally similar in each environment. The main difference was that in the part-task experiment, the first officer role was played by a surrogate assuming a passive first officer role.

In addition to the aircrew, it was also necessary to provide models for the approach, tower, and ground controllers that guided the aircrew through the approach and landing. With these players and their procedures in place, the next step was to reproduce the results of the part-task scenarios. Given the three display conditions and the four approach event conditions, there were twelve possible scenarios. Of the twelve possible scenarios, ten were selected for use in the part-task experiment. The conditions for the two remaining scenarios would have produced non-meaningful scenarios (e.g. terrain misalignment in VMC) and were not run. To date, eight of the scenarios have been executed in the human performance modeling environment. For the purposes of this study, we focused on the results from two of the part-task scenarios: the scenario in which the captain used the baseline instrument displays during the IMC approach and the scenario in which the SVS display was also available during the nominal IMC approach. In the model runs, aircrew procedures closely followed those of the skilled aircrew subjects in the part-task experiments and yielded eye-tracking data that was consistent with the human results from the part-task scenarios.

With the part-task experiment data qualitatively reproduced by the models, the next step was to develop a model for an Enhanced-SVS flight deck configuration. The SVS replicated most of the functionality of the PFD. We wanted to provide an Enhanced-SVS that subsumed the functionality of the PFD so that it could be used as the sole flight deck navigation instrument. To accomplish this, the flight mode annunciators, the functionalities missing from the standard SVS display, were added to the display. With the functionality of the PFD now fully duplicated by the Enhanced-SVS, the PFD was removed from the configuration. At this point, we were ready to complete the model runs that had not been done as part of the part-task experiment—the captain's use of the Enhanced-SVS flight deck during a nominal IMC approach.

RESULTS

Table 1 provides eye-tracking data for a representative flight segment for the human subject and model trials. The human subject data was provided by NASA (Goodman et al., 2003). The model data was derived from the D-OMAR simulation trials. The columns present the percentage of dwell time that each subject, human or model, devoted to the out-the-window view (OTW), to each of the principal flight deck instruments, and to "other" identified areas in their field of view. For the human subjects, the row labeled "off" accounts for the percentage of dwell time for which the eye cursor was centered on an undefined area or for which the data was invalid (e.g., subject blinks) (Goodman et al., 2003). The data in the table covers the flight segment between the initial approach fix (IAF) and the final approach fix (FAF) on the approach to Santa Barbara Municipal airport runway 33 left. The approach plate provided to the subjects for the NASA part-task experiment is shown in Figure 1. The IAF and FAF are identified in the approach plate. Columns 2 through 4 provide data from the baseline-IMC approach for the three human subjects. Column 5 provides data from the baseline-IMC approach for the D-OMAR model. Columns 6 through 8 provide data from the SVS-Equipped IMC approach for the three human subjects. Column 9 provides data from SVS-Equipped IMC approach for the D-OMAR model. Column 10 provides model data for the IMC scenario using the Enhanced-SVS attitude display that made it possible to eliminate the PFD.

Table 1 Human Subject and Model Eye-tracking Data (Initial Approach Fix to Final Approach Fix)

	Subject 3	**Subject 4**	**Subject 5**	**Model**		**Subject 3**	**Subject 4**	**Subject 5**	**Model**		**Model**
Condition	IMC	IMC	IMC	IMC		SVS	SVS	SVS	SVS		Enh-SVS
off	2.24	3.90	5.15			2.40	4.76	1.53			
OTW	0	1.00	0.35	6.46		0	4.82	0.06	6.07		6.52
SVS						**24.02**	**15.25**	**11.84**	**26.98**		**37.28**
PFD	**43.63**	**55.95**	**34.42**	**37.65**		**27.93**	**31.36**	**42.57**	**26.81**		
NAV	**32.62**	**33.52**	**54.71**	**48.24**		**28.96**	**39.55**	**37.39**	**34.46**		**48.18**
other	21.51	5.49	5.36			16.69	4.26	6.62			

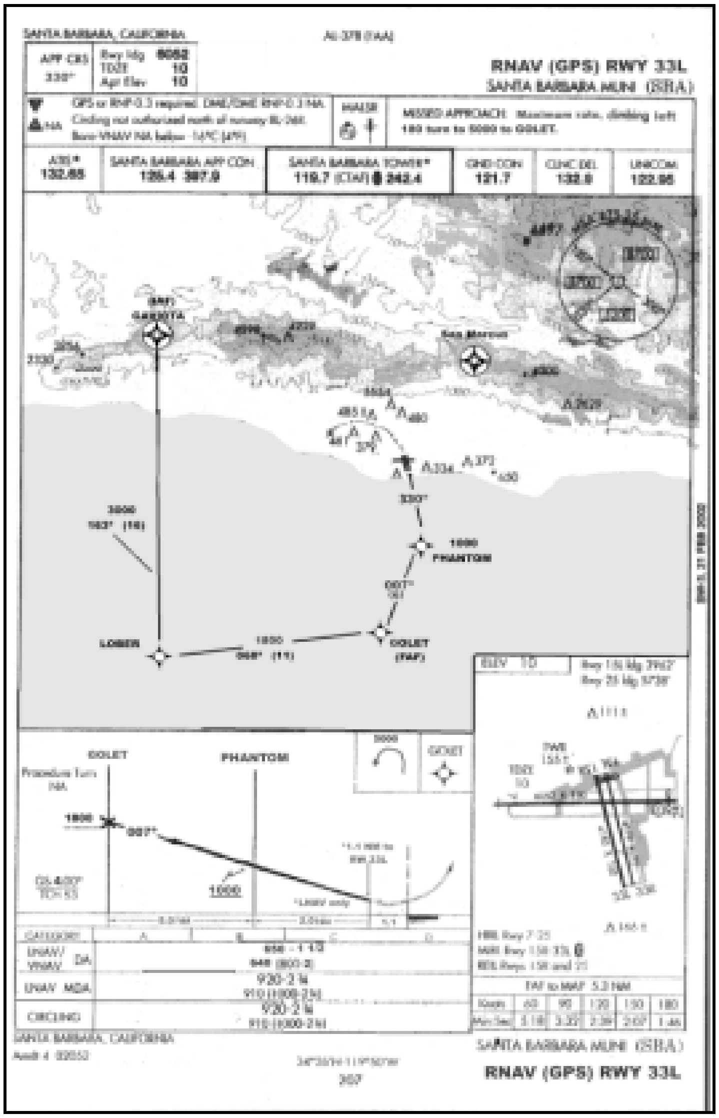

Figure 1 Approach Plate for SBA Runway 33 Left

The baseline-IMC scenarios included a flight deck with a single attitude display (the PFD) and a single navigation display (the HSI). In the SVS scenarios, the SVS (provided only to the captain) served as a second attitude display. In the part-task experiment with commercial pilots as subjects, the presence of the SVS had the effect, across the experiments and for three of the four flight phases within the experiments, of shifting the time split between the attitude display(s) and the navigation display in direction of the attitude displays. The data in the table readily demonstrates the reduced time devoted to the navigation display (the NAV row in the table) by each subject in the SVS-equipped configuration. The procedures adopted for the aircrew models when using the baseline flight deck and the SVS-equipped flight deck that included a scan of the SVS, lead to aircrew model behaviors similar to those for the human subject experiments. With the SVS-equipped flight deck, the modeled captain devoted more time to the attitude displays and less time to the navigation display when executing the same tasks.

We have run each of the four SVS scenarios using the Enhanced-SVS: the nominal approach IMC condition; the late reassignment IMC condition; the missed approach IMC condition; and the terrain mismatch IMC condition. Table 1, representing one flight segment (IAF to FAF) from the nominal IMC condition, demonstrates that when using a single attitude display, as anticipated, the allocation of time devoted to the attitude display and the navigation display reverted to that seen in the baseline configuration. The balance varies through the phases of the scenarios, particularly around the maneuvers at flight path waypoints, but is once again consistent with the pattern in the baseline condition.

DISCUSSION

An important observation from the part-task scenario study was that when the flight deck was augmented with the SVS display, the balance of scan time allocated to the attitude displays increased at the expense of time devoted to the navigation display. This was true even though the underlying tasks to be accomplished remained the same. The addition of the new instrument altered well-established pilot scan patterns and raised the concern that the two-attitude-displays configuration (PFD + SVS) at least encouraged and perhaps required more scan time to accomplish the same attitude-related functions.

Given the scan time allocation observation, our attention turned to investigating a means to retain the advantages provided by the SVS, while mitigating the effect of having a second flight deck attitude display. An Enhanced-SVS that combined the functions of the primary flight display and the synthetic vision system in a single attitude instrument was readily adapted as a potential solution. The modeling framework provided a means to explore the possibility that this option would achieve the desired effect.

Using the D-OMAR model that had been validated to point of demonstrating that it produced similar pilot scanning data for both the baseline and SVS conditions, it was possible to simulate and then evaluate the pilot's behavior using the easily modeled Enhanced-SVS as the single attitude display. Exercising this model produced the anticipate shift back to the more efficient scanning behavior observed in the baseline condition while preserving the advantages of the synthetic vision system capabilities.

These results are important not only for the substantive predictions they generated, but also as a concrete example of the ways in which human performance models can contribute to design decisions early in the design process. The cost in time and effort of exercising the model condition in order to demonstrate the validity of the hypothesis that the Enhanced-SVS would be a more efficient configuration was a very small fraction of the cost of building a revised human-in-the-loop simulation, running a new set of subjects, and analyzing the resulting eye-movement data.

ACKNOWLEDGMENTS

The research reported on in this paper was supported by the NASA Ames Research Center, Aviation Safety Program, System-Wide Accident Prevention project, Human Performance Modeling element led by Dr. David Foyle.

REFERENCES

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.

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. (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.

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

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

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

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.