Small Aircraft Transportation System Higher Volume Operations Flight Experiment

Daniel M. Williams,* Jennifer L. Murdoch,† and Catherine H. Adams‡
NASA Langley Research Center, Hampton, Virginia 23681

DOI: 10.2514/1.20403

This paper summarizes findings from the Small Aircraft Transportation System Higher Volume Operations Flight Experiment. The higher volume operations concept improves efficiency at nontowered, nonradar airports in instrument meteorological conditions. The success of the higher volume operations concept is based on pilot acceptability as determined through objective and subjective assessments when compared with the procedural control operations in use today at nontowered, nonradar-controlled airfields in instrument meteorological conditions. Flight experiment data indicate that the concept is viable. The experiment, flown on a general aviation aircraft, used a subset of the Higher Volume Operations Simulation Experiment scenarios and evaluation pilots to validate the simulation experiment results. Results reveal that all 12 low-time instrument-rated pilots preferred Small Aircraft Transportation System Higher Volume Operations when compared with current procedural separation operations. These pilots also flew the higher volume operations procedures safely and proficiently without additional workload in comparison to today's system. Detailed results of pilot flight technical error and their subjective assessments of workload and situation awareness are presented.

I. Introduction

IN the United States, the current National Airspace System of hub-and-spoke operations has served its purpose well, but it is beginning to reach a capacity plateau. Because of increasing demand on the system and with only modest potential gains in the number of flights, the system could reach gridlock within the next 10–15 years [1,2]. Several new, small, efficient aircraft are being developed by Honda, Cessna, Diamond, Eclipse, Safire, Cirrus, Lancair, Adam Aircraft, and others to provide point-to-point service and make use of small airports, many without control towers that lie outside air traffic control (ATC) radar coverage.

Wheninstrument meteorological conditions restrict operations to instrument flight rules at nontowered, nonradar airports, ATC uses procedural separation that restricts operations to only one approaching or departing aircraft at a time: the "one-in/one-out" paradigm that severely limits the operational throughput at these airports. Air charter operators might be compelled to use these airfields if the operational efficiency can be improved. The Small Aircraft Transportation System (SATS) project breaks the one-in/one-out paradigm and expands capacity by allowing multiple, simultaneous operations. The concept of operations that achieves this goal is termed SATS higher volume operations (HVO) [3].

Several human-in-the-loop experiments were conducted with low-time instrument-rated pilots to address if these pilots could fly SATS HVO safely, and proficiently, and with acceptable workload and situation awareness. And based on a pilot acceptability of SATS HVO through these experiments, pilots will be able to fly SATS HVO with similar, if not lower workload, flight technical error, and better situation and navigation awareness when compared with the one-in/one-out procedures in use today. Key to this HVO validation process were the results from the HVO flight experiment described here that validate the HVO concept for normal conditions [3].

Presented as Paper 7421 at the AIAA 5th ATIO and 16th Lighter-Than-Air Sys Tech. & Balloon Systems Conferences, Arlington, Virginia, 26–28 September 2005; received 6 October 2005; revision received 4 January 2006; accepted for publication 9 January 2006.

*Research Engineer, Aviation Operations and Evaluation Branch, Mail Stop 152. Member AIAA.

†Research Psychologist, Aviation Operations and Evaluation Branch, Mail Stop 156A.

‡Research Psychologist, Crew Systems Branch, Mail Stop 152.

II. Background

A. SATS HVO Overview

The SATS HVO concept is based on a distributed decision-making environment that assumes major decision-making responsibility be held by the pilot, and resource mitigation be held by ground-based automation, the airport management module (AMM). The concept uses a newly defined flight operations area called a self-controlled area (SCA), established during instrument meteorological conditions around SATS designated airports (i.e., nontowered, nonradar airports). Within the SCA, pilots, using advanced airborne systems, would have the ability and responsibility to maintain separation between themselves and other similarly equipped airplanes. Aircraft operating in this airspace would need special avionics, for example, automatic dependent surveillance-broadcast (ADS-B), a two-way data link, and appropriate self-separation tools in order to participate. The AMM provides appropriate sequencing information to the arriving aircraft. The AMM distributes an arrival sequence and broadcasts the total number of arriving aircraft in the SCA. It does not, however, provide separation, altitude assignments, or sequence departures.

Aircraft will approach a SATS airport on an instrument flight rules clearance granted by ATC to a transition fix above the SCA. This fix is also an initial approach fix (IAF) for an instrument approach procedure.§Before reaching the fix, the pilot requests a landing assignment from the AMM through their onboard system. The AMM responds with the SCA entry procedure (standby, vertical, or lateral), relative sequence information (follow <Callsign>), and missed approach hold fix assignment (e.g., Annie or Cathy). The AMM only sequences arrivals (including missed approach aircraft), not departures. Nominally, up to four arriving aircraft are allowed in the SCA before entry is denied (AMM issues a standby), though this constraint can be affected by local airspace restrictions.‖ Following their entry assignments and the HVO procedure to "descend to lowest available altitude," pilots are deconflicted from other arriving aircraft (i.e., the AMM reserves a slot at one of the IAFs for each SCA aircraft until it lands or departs the SCA).

§GPS-T instrument approach procedures were chosen as a basis for this concept, although other instrument approach procedures could be used.

‖The number of arriving aircraft, including those executing a missed approach, are limited by the holding altitudes available for the approach. Figure 1 shows the nominal approach design with four potential holding segments.

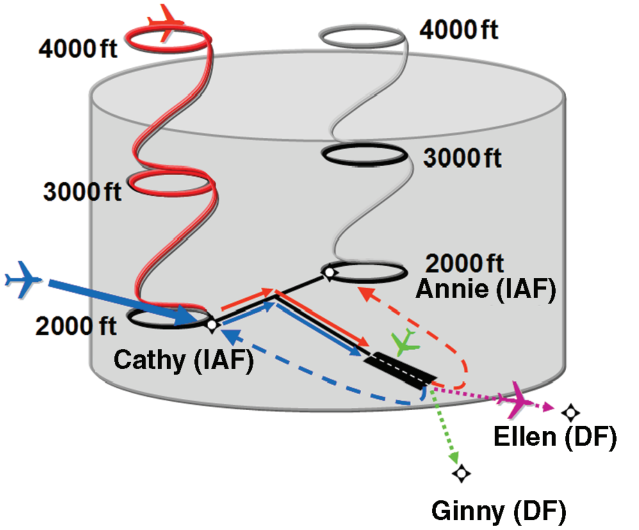

Fig. 1 SATS HVO example.

Many of the features of the SATS HVO concept are depicted in Fig. 1. SATS arrivals (red and blue aircraft) to the IAFs with alternating missed approaches, and departures (green and purple aircraft) to the departure fixes (DFs):

1) Blue depicts entering the SCA, having coordinated descent with ATC when no other aircraft were assigned to Cathy; the missed approach is a blue dashed path, and the AMM returned: "Lateral entry, follow none, missed approach Cathy."

2) Red depicts having arrived by ATC instruction to the transition fix above the SCA at 4000 ft with one other Cathy assignment, and the AMM returned: "Vertical entry, follow blue, missed approach Annie."

3) Purple depicts departing the SCA via the departure procedure and contacting ATC before a DF.

4) Green depicts having been released by ATC to depart (within the departure window), holding short, and using onboard tools to find an open slot in the arrival stream to take the runway and depart.

Pilots given a standby sequence can track the number of aircraft in the SCA to estimate their delay as they continue to their clearance limit, the transition fix at an altitude above the SCA, and hold. When the pilot receives an AMM entry message with sequence and missed approach information, the pilot checks for an available holding altitude, and will request descent from ATC. The pilot can then determine if further descent is prudent by following the lowest available altitude procedure at the IAF (clearing for traffic below is the pilot's self-separation responsibility in the SCA). Pilots initiate their approach once adequate spacing behind the lead aircraft has been met (determined through either a generic rule-based spacing procedure, that is, safe for all combinations of aircraft performance, or by using an onboard self-spacing tool). The AMM reserves a holding slot for assigned missed approaches. A pilot executing a missed approach would climb to the lowest available altitude at their assigned missed approach holding fix and would then be sent a new arrival sequence.

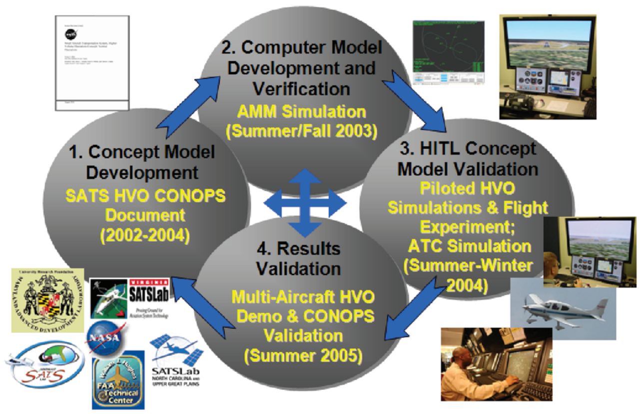

Fig. 2 HVO validation.

For SATS departures, pilots will file flight plans with a SATS departure procedure to a departure fix, that is, Fig. 1 Ellen or Ginny. Just as in today's nonradar environment, the pilot should expect a clearance void time and potentially a release time restriction as part of their instrument flight rules clearance. This affords seamless integration with today's instrument flight operations. Within this ATC departure window, they will use onboard information/tools to deconflict themselves with landing traffic, for example, to ensure no arriving aircraft within 5 n mile of the airport. The pilot would then depart and contact ATC according to the departure procedure before entering ATC controlled airspace.

B. Validating the SATS HVO Concept and the Use of Piloted Studies

The SATS HVO research team developed and applied a four-phase process to design and validate the HVO concept, depicted in Fig. 2.

Phase 1 involved HVO concept development. The key safety properties of a draft HVO concept were established by mathematical verification methods based on formal logic and theorem proving [4]. This initial study began formally verifying that self-separation can be maintained when pilots adhere to the HVO procedures (including AMM logic).

Phase 2 involved the development of computer simulations. The AMM function and associated algorithms were verified and validated using a representative set of normal HVO scenarios, that is, flown to procedure without deviations. Initial batch simulation estimates were also conducted that showed a four-fold throughput improvement by HVO when compared with the one-in/one-out environment [5].

Phase 3, providing the primary material of this paper, determined if HVO functionality and procedures were safe, flyable, and acceptable to pilots. Human-in-the-loop scenarios were developed that compared SATS HVO to the one-in/one-out procedural control environment available today (baseline).

Low-time instrument-rated evaluation pilots provided experimental data and subjective feedback as they flew the scenarios in experiments using progressively higher fidelity simulation, from a medium fidelity general aviation computer simulation [5,6] to the Cirrus SR22 aircraft in flight. Consistent early results across the various experiment platforms, including the high-fidelity HVO flight experiment, provided confidence in the simulation environment, and so later studies of the HVO concept were conducted with the simulation platform [7]. Subsequently, two additional simulation experiments determined if nonnormal procedures were acceptable to the pilot as well as the procedure support automation developed by NASA researchers [8,9]. Also, an ATC simulation study was completed and focused on determining controller acceptability of the concept model [10].

Phase 4 was a proof-of-concept public demonstration of six different aircraft that flew the SATS HVO procedures in the 2005 SATS Technology Demonstration held in Danville, Virginia. All four phases provided feedback to the improvement of SATS HVO and ultimately toward recommending a viable way to improve upon the one-in/one-out procedure in place in the National Airspace System today.

C. The Need for the HVO Flight Experiment

Because researchers used an iterative process to increase HVO technology readiness level through research and development, the flight experiment was a critical step to validate pilot acceptability of a high-fidelity HVO system in an actual flight environment. An analysis of flight technical error from the HVO Flight Experiment also validated the HVO Simulation Experiment and, subsequently, the simulation environment (the SATS Air Traffic Operations Lab) as a platform worthy of conducting further detailed HVO analyses [7]. Also, developing and implementing HVO systems and procedures for use in the Cirrus SR22 became a lead-in for a strong

NASA leadership of the phase 4, full-architecture, multiaircraft, proof-of-concept demonstration.

III. Experiment Description

A. HVO Validation Experiment: Objectives and Methodology

Determining pilot acceptability of HVO meant comparing SATS HVO to the one-in/one-out procedural control environment available today (baseline) and answering two research objectives:

1) Can pilots safely and proficiently fly the airplane while performing SATS HVO procedures?

2) Do pilots perceive that workload, while using HVO procedures and tools, is no greater than flying in today's system?

The following hypotheses were developed to meet the research objectives:

1) Flight technical error (baseline and SATS scenarios): subject pilots will fly within the FAA practical test standards for the instrument rating 100% of the time during all scenarios. Deviations from assigned flight paths [i.e., rms error (RMSE) values will be equivalent across all scenarios.

2) Subjective workload (baseline and SATS scenarios): equivalent workload ratings will be associated with all scenarios.

3) Situation awareness (baseline and SATS scenarios): equivalent situation awareness ratings will be associated with all scenarios.

B. Experimental Design

The experiment design matrix shown in Fig. 3 includes the five scenarios flown in the HVO Simulation Experiment [6]. Three of these scenarios were repeated in the flight experiment. For the simulation experiment, the experiment design used for data collection was a 2 (procedure type) ×5 (scenario type), within-subject design in which the same 15 participants (i.e., low-time instrument-rated pilots) were assigned to each experimental cell (i.e., test condition). Simulation evaluation pilots 1 through 15 (S_{1-15}) were asked to perform all 10 test conditions in partially counterbalanced order under simulated instrument meteorological conditions. For the flight experiment, a 2 (procedure type) ×3 (scenario type), within-subject design was used for data collection and 12 of the same 15 evaluation pilots (F_{1-12}) performed six test conditions twice in partially counterbalanced order under simulated instrument meteorological conditions. Dependent measures included pilot flight technical error and subjective assessments of workload and situation awareness.

		PROCEDURE TYPE	
		Baseline	SATS
SCENARIO TYPE	Departure	S_{1-15}	S_{1-15}
	Approach Without Traffic	$S_{1-15} F_{1-12}$	$S_{1-15} F_{1-12}$
	Approach With Virtual Traffic	$S_{1-15} F_{1-12}$	$S_{1-15} F_{1-12}$
	Approach to Missed Approach with Virtual Traffic	$S_{1-15} F_{1-12}$	$S_{1-15} F_{1-12}$
	Approach with Piloted Traffic (Linked Simulation)	S_{1-15}	F_{1-15}

S=Simulation Subjects
F= Flight Subjects

Fig. 3 HVO validation: experiment design matrix.

C. Scenarios and Profiles

Baseline and SATS procedures were conducted for each scenario. During baseline procedures, ATC allowed only one aircraft on the GPS nonprecision approach at a time and provided clearances throughout the scenario, for example, holding, descents, and initiating the approach. The SATS procedures were also based off of a GPS nonprecision approach procedure and gave the pilot the responsibility to follow the AMM designated sequence, maintain separation from other SCA aircraft (including descent to the lowest available altitude), initiate the approach, and self-space from a lead aircraft. The experiments were designed to examine the HVO concept through a representative set of sample scenarios:

1) Departure with approaching traffic. This scenario was flown in the simulation experiment only (not reported in this paper) [6].

2) Approach without traffic (no holding required). The pilot task was to descend and fly the approach via ATC clearance during the baseline scenarios. During the SATS scenarios, the pilot requested and received a "lateral entry, follow none" sequence from the AMM, was handed off into the SCA from ATC, and self-initiated the approach. Typically the durations of the SATS scenarios were equivalent to those of the baseline scenarios.

3) Approach with virtual traffic (holding required). This scenario clearly differentiated baseline from HVO efficiency. In baseline, the pilot waited behind two other aircraft in holding until they had landed and ATC provided clearance to begin the approach (i.e., 30+ minutes in holding). In HVO, the pilot followed the AMM sequencing behind the two other aircraft, self-separated from the other aircraft in the SCA, and self-initiated the approach by following advisories provided by a self-spacing software tool. The SATS scenarios' durations were about half that of the Baseline scenarios.

4) Approach to missed approach with virtual traffic (holding required). This scenario included having the pilot fly the missed approach. SATS scenarios required flying the missed approach while

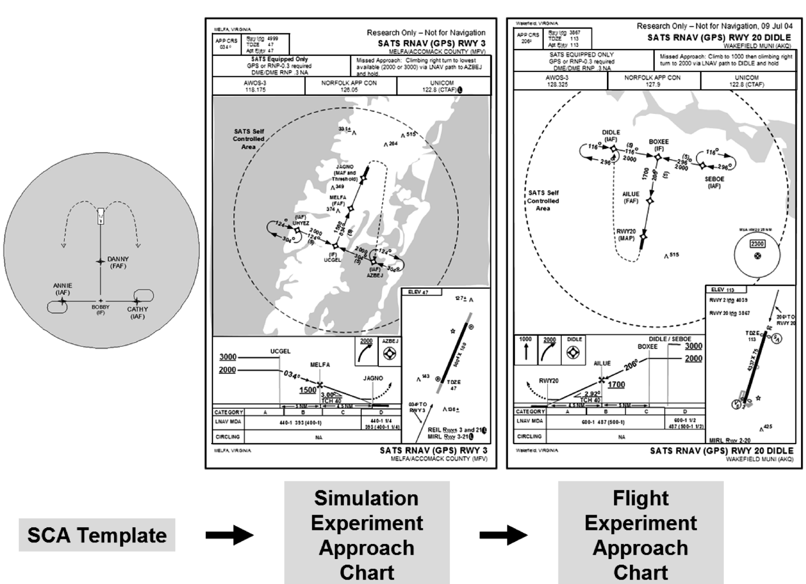

Fig. 4 HVO procedure development progression: concept to flight experiment.

self-separating in the SCA. The SATS scenarios' durations were about two-thirds as long as the baseline scenarios' durations.

5) Multipilot linked simulation approaches (holding required). This scenario was flown with four linked simulators (not reported in this paper) [6].

The experimental scenario flight profiles were based on GPS-T instrument approach procedures. Figure 4 depicts the progressive development of the experiment flight profile that began with the template SCA in the HVO concept of operations document and was implemented as shown in the simulation and flight approach charts. Evaluation pilots used these to fly experiment scenarios (baseline charts not shown). The approach chart is a key element of the pilot procedures and coupled with task training, all evaluation pilots were able to fly the HVO procedures (described earlier with Fig. 1) and the accompanying nonprecision GPS-T instrument flight approach procedure.

D. Progressing from Simulation to Flight: Pilot Criteria, Training, and Apparatus

In choosing a subject pool for the validation experiments, researchers identified criteria for evaluation pilots who were capable, but not overly experienced. Although not tested, it was postulated that if the low-time instrument-rated subject pool validated the concept positively, then more experienced pilots would do the same. Subject pilot criteria included low-time (less than 1000 flight hours), an instrument rating, and legal currency to fly in instrument flight rules. None of the participants had previously flown a Cirrus SR22 aircraft, worked as a flight-crew member for an air carrier within the last year, or flown for the military.

Twelve subject pilots were randomly drawn from the original 15 evaluation pilots that participated in the HVO Simulation Experiment. This reduced training requirements for the HVO Flight Experiment and allowed pilots to progress logically from hand-flying a medium-fidelity general aviation simulator to the Cirrus SR22 aircraft. In the post experiment debriefing, all 12 pilots were comfortable that the preexperiment training they received was adequate to fly experimental scenarios.

Figure 5 shows the general aviation simulator and the Cirrus SR22's instrument panel used for the experiments. Common software across platforms was developed to drive the multifunction display (MFD) and the horizontal situation indicator (HSI) displays. Variation between the simulation and flight profiles was deliberately minimized so as not to alter the experiment objectives or hypotheses. Pilots were tasked to manually fly the scenarios in simulated instrument meteorological conditions using traditional round-dial instruments for primary flight guidance information (i.e., without autopilot). Pilots were tasked to meet FAA instrument rating practical test standards criteria during all scenarios (both SATS and baseline) [11]. Another research postulate not tested was that if pilots performed HVO tasks well with the simple NASA Cirrus SR22 research avionics configuration, then better avionics (e.g., a primary flight display) might simplify tasks even further.

Figure 6 shows the NASA Cirrus SR22 used in the HVO Flight Experiment. On board were the research software components that enabled evaluation pilots to fly a subset of the HVO simulation scenarios. The research MFD and HSI were the primary interface devices that pilots used while they hand-flew the SR22 through the baseline and SATS procedures. The HSI provided course and course-deviation information that corresponded to the same GPS nonprecision approach procedure being flown. The MFD had a moving map with the current position, current leg, the navigation path, and virtual scenario traffic depicted (a simulated automatic dependent surveillance-broadcast environment). For SATS procedures, the MFD also provided HVO procedure support through the "pilot advisor," which advised pilots sequentially through HVO rules and procedures. For SATS procedures, self-spacing software provided support for the initiation of approaches. The MFD contained SATS conflict detection and alerting capability, but only normal HVO scenarios (i.e., without conflicts), were flown. For both baseline and SATS procedures, simulated virtual aircraft and ATC audio communications were fed into the pilot's headset to build a realistic scenario environment. These research systems were on board for the flight experiment and were an interim configuration that formed the foundation for the phase 4 full-architecture, multiaircraft, proof-of-concept demonstration.

A large screen capture of the SATS research MFD is shown in Fig. 7. The information presented on the MFD was of significant concern in the workload and situation awareness evaluations. The HVO Simulation Experiment also reported usability results of the interface. Baseline procedures did not include the HVO specific information conveyed through an alert window (shown with the Advisor), nor the AMM information window (upper right, middle), lead aircraft double-chevron highlighting, or pilot advisor window (shown with "Open Approach").

Fig. 5 HVO experiment platform progression (note commonality of HSI and MFD).

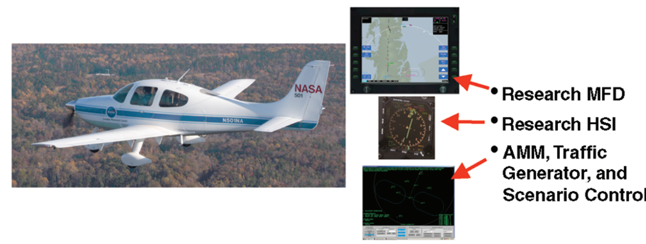

Fig. 6 SATS HVO Cirrus SR22 with flight experiment apparatus.

Fig. 7 Research MFD (note: leading aircraft, AMM message; Pilot Advisor "OPEN APPROACH").

IV. Analysis of Results

Data were collected both during and at the end of the scenarios as well as the end of each evaluation pilot's session. Quantitative results to follow include flight technical error of airspeed, altitude, and lateral path deviation. Qualitative results shown are from workload assessments and situation awareness assessments including traffic and navigation guidance awareness.

A. Flight Technical Error

1. *RMSE with Respect to Airspeed Target on the Approach*

Pilots were instructed to fly the same target airspeeds for both the SATS and baseline scenarios. Target airspeeds were chosen for the three segments of the GPS-T nonprecision instrument approach procedure: initial (initial approach fix to intermediate fix), intermediate (intermediate fix to final approach fix), and final (final approach fix to missed approach point). The target airspeeds for the initial, intermediate, and final segments were 120, 110, and 100 kn, respectively, for the experiment. Instrument rating practical test standards require the pilot maintain $+/-10$ kn of the target airspeed [11]. Data were analyzed by way of repeated analysis of variance (ANOVA) tests, with the main effect of procedure type being of primary interest, and a 5% significance level (i.e., $p \leq 0.05$) for the statistical analyses was set a priori [12,13].

A significant difference was found to exist between the airspeed target RMSE values associated with each procedure type ($p = 0.003$). A RMSE airspeed deviation of 5.82 kn occurred during the baseline scenarios ($M = 5.82$, $SD = 1.81$, and $N = 72$, where M is the mean, SD is the standard deviation, and N is the sample number, respectively), and a RMSE airspeed deviation of 4.82 kn occurred during the SATS scenarios ($M = 4.82$, $SD = 1.29$, and $N = 72$).

The hypothesis that evaluation pilots would "fly within the FAA practical test standards for the instrument rating 100% of the time during all scenarios" was supported by the airspeed target RMSE values[¶] as was the hypothesis that "deviations would be at least equivalent across all scenarios." Fig. 8 indicates that evaluation pilots maintained airspeed with respect to an assigned target value more accurately when they performed the SATS scenarios than when they performed the baseline scenarios. Based on the limited context of these findings, pilots can fly SATS HVO procedures safely and proficiently when compared with today's system (baseline). Note that the slightly better overall performance during SATS procedures may reflect the pilots reacting to "too fast" and "too slow" procedure support advisories that were implemented with a $+/-10$ kn window about the procedure's target airspeeds.

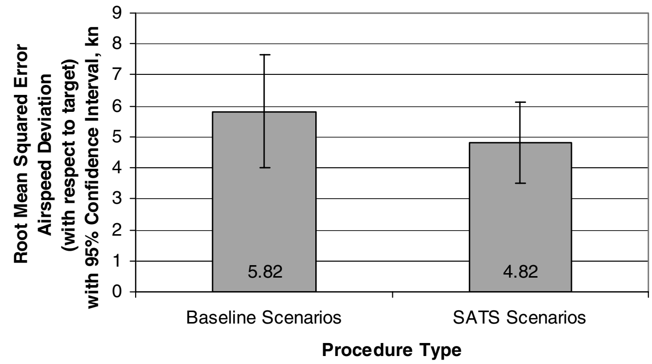

Fig. 8 Flight experiment deviation from airspeed target.

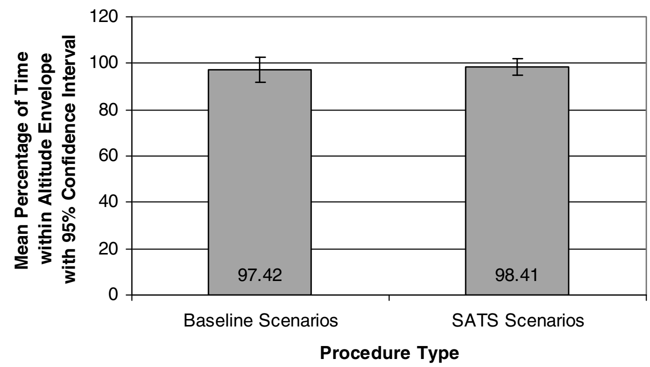

Fig. 9 Flight experiment percent time in altitude envelope.

2. *Percent Time Within Altitude Envelope on the Approach*

Pilots were instructed to fly within the same practical test standards altitude envelope for both the baseline and SATS scenarios: -100 ft of "at or above" altitudes, and $+100$ and -0 ft for the minimum descent altitude (MDA) until the missed approach point or visual transition to landing [11].

No significant difference was found to exist between the percentages of time within the altitude envelope associated with each procedure type ($p = 0.300$). During the baseline scenarios, evaluation pilots flew within the defined altitude envelope 97.42% of the time ($M = 97.42$, $SD = 5.44$, and $N = 72$). During the SATS scenarios, evaluation pilots flew within the defined altitude envelope 98.41% of the time ($M = 98.41$, $SD = 3.58$, $N = 72$).

The hypothesis that subject pilots would "fly within the FAA's practical test standards for the instrument rating 100% of the time during all scenarios" was not supported, but this does not diminish the acceptability of the SATS HVO concept, because the hypothesis that "deviations would be equivalent across all scenarios" was supported. Figure 9 indicates that evaluation pilots maintained altitude within an assigned envelope equally well when they performed the baseline scenarios and when they performed the SATS scenarios. Based on the limited context of these findings, pilots can fly SATS HVO procedures safely and proficiently when compared with today's system (baseline).

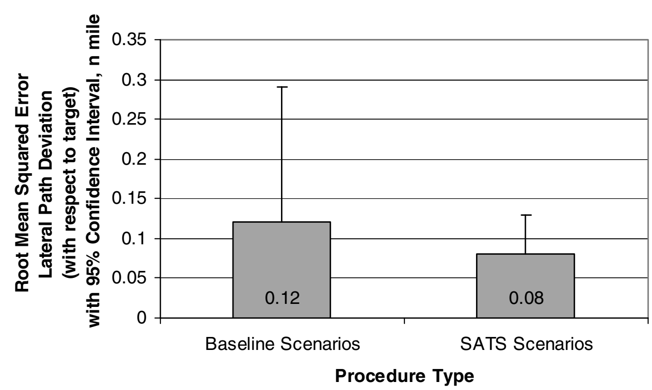

Fig. 10 Flight experiment deviation from lateral path target.

3. *RMSE with Respect to Lateral Path Target on the Approach*

Pilots were instructed to use the target path as the approach flight path during both the baseline and SATS scenarios. Instrument practical test standards are to maintain within 3/4-scale deflection of the course-deviation indicator on the research HSI [11]. Full-scale deflection represented 0.3 n mile through the approach procedure.

In the flight experiment, a significant difference was found to exist between the lateral path deviation target RMSE values associated with each procedure type ($p = 0.045$). A RMSE lateral path deviation of 0.12 n mile occurred during the baseline scenarios ($M = 0.12$, $SD = 0.17$, and $N = 72$), and a RMSE lateral path deviation of 0.08 n mile occurred during the SATS scenarios ($M = 0.08$, $SD = 0.05$, and $N = 72$).

The hypothesis that pilots would "fly within the FAA's practical test standards for the instrument rating 100% of the time during all scenarios" was supported as was the hypothesis that "deviations would be at least equivalent across all scenarios." Fig. 10 indicates that evaluation pilots maintained lateral path deviation with respect to an assigned target value more accurately when they performed the SATS scenarios than when they performed the baseline scenarios. Based on the limited context of these findings, pilots can fly SATS

[¶]Statistical confidence was 95% (standard practice) that RMSE deviation data shown are accurate.

HVO procedures safely and proficiently when compared with today's system (baseline). Note that the improved pilot performance in the SATS scenarios was unexpected because the HSI guidance given to the pilots was the same for baseline and SATS procedures. It can be surmised that the SATS pilot is more "engaged in the process," better aware of where the other aircraft in the arrival sequence are, and flying with better precision by anticipating the next turn and descent, vs. reacting to an ATC clearance. Although observed, this was not studied specifically, and is worthy of further exploration.

B. Subjective Assessments of Workload

Evaluation pilots used the modified Cooper–Harper rating scale to rate the level of workload that they experienced during each of the experiment's six test conditions. Workload ratings could range on a scale from 1 (i.e., the instructed task was very easy/highly desirable, operator mental effort was minimal, and the desired performance was easily attainable) to 10 (i.e., the instructed task was impossible; it could not be accomplished reliably) [14]. Because evaluation pilots performed each of the 6 test conditions twice, each provided 12 workload ratings. For each evaluation pilot, the two workload ratings associated with a given test condition were averaged together to yield a set of six workload ratings. As reported next, nonparametric tests were employed as a conservative method for analyzing workload ratings associated with discrete rating scale items.

Evaluation pilots reported experiencing the following workload ratings when they performed different types of scenarios using baseline and SATS procedures ($N = 12$ for each scenario):

1) Scenario 3 (baseline, approach without traffic), $M = 2.08$, $SD = 0.51$

2) Scenario 4 (SATS, approach without traffic), $M = 1.46$, $SD = 0.45$

3) Scenario 5 (baseline, approach with virtual traffic), $M = 2.67$, $SD = 0.69$

4) Scenario 6 (SATS, approach with virtual traffic), $M = 1.58$, $SD = 0.42$

5) Scenario 7 (baseline, approach to missed approach with virtual traffic), $M = 2.29$, $SD = 0.54$

6) Scenario 8 (SATS, approach to missed approach with virtual traffic), $M = 1.67$, $SD = 0.44$

A series of Wilcoxon tests [15] was performed to determine if evaluation pilots reported experiencing different levels of workload when different types of scenarios were performed using different types of procedures. Because the differences between the workload ratings associated with particular pairs of scenarios were of primary interest (i.e., scenarios 3 vs. 4, 5 vs. 6, and 7 vs. 8), only the results of the Wilcoxon tests associated with these scenario pairings are discussed here.

As shown by the mean workload ratings plotted in Fig. 11, evaluation pilots reported experiencing higher levels of workload

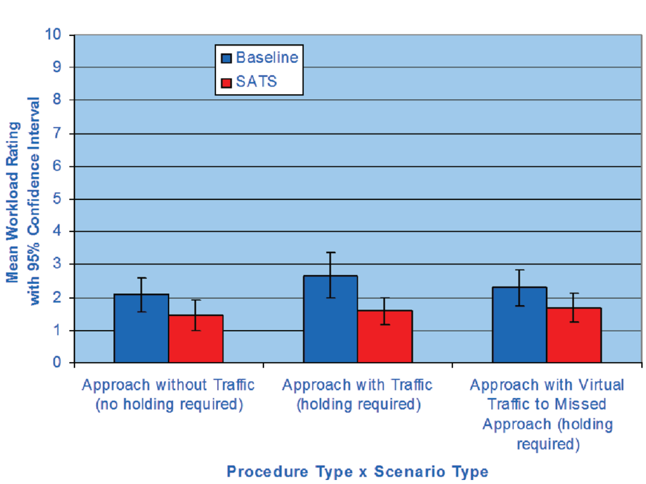

Fig. 11 Mean workload ratings associated with procedure type × scenario type.

when scenarios were performed using the baseline procedures than when scenarios were performed using the SATS procedures.

The Wilcoxon tests revealed that, at a statistically significant level, evaluation pilots reported experiencing higher levels of workload when performing scenario 3 as compared with scenario 4 ($p = 0.0357$) and when performing scenario 5 as compared with scenario 6 ($p = 0.0049$). Statistically, evaluation pilots reported experiencing equivalent levels of workload when performing scenario 7 as compared with scenario 8 ($p = 0.0881$). In summary, these data illustrate that pilots reported their workload decreased when they flew SATS HVO scenarios (surpassing the experimental hypothesis that "equivalent workload ratings will be associated with all scenarios").

C. Subjective Assessments of Situation Awareness

Evaluation pilots were administered a situational awareness rating technique (SART) instrument that included the three dimensions of demand, supply, and understanding, as well as two independent dimensions of traffic awareness and navigation guidance [16]. For the dimensions of demand, supply, and understanding, evaluation pilots used a scale ranging from 1 (low) to 7 (high) to rate each dimension. The formula shown next was used by the experimenters to calculate global SART ratings:

$$SA = \text{Understanding} - (\text{Demand} - \text{Supply})$$

Global SART ratings can range from 1 [representing a low level of situation awareness (SA)] to 14 (representing a high level of SA). In the current study, calculated SART ratings ranged from 1 to 11.5. For traffic awareness, scores ranging from 2 to 7 on a scale of 1 (low) to 7 (high) were collected from the evaluation pilots, and scores ranging from 1.5 to 7 on a scale of 1 (low) to 7 (high) were collected for navigation guidance SA. As reported next, Wilcoxon tests were used to analyze the evaluation pilots' SART scores, traffic awareness scores, and navigation guidance SA scores. Wilcoxon tests (i.e., nonparametric within-subject tests appropriate for analyzing two related samples of ordinal data) were employed as a conservative method for analyzing SA ratings associated with discrete rating scale items.

Evaluation pilots reported the following SART ratings when they performed different types of scenarios using baseline and SATS procedures ($N = 12$ for each scenario):

1) Scenario 3 (baseline, approach without traffic), $M = 6.92$, $SD = 1.96$

2) Scenario 4 (SATS, approach without traffic), $M = 8.08$, $SD = 1.89$

3) Scenario 5 (baseline, approach with virtual traffic), $M = 5.38$, $SD = 2.76$

4) Scenario 6 (SATS, approach with virtual traffic), $M = 7.75$, $SD = 1.74$

5) Scenario 7 (baseline, approach to missed approach with virtual traffic), $M = 5.79$, $SD = 2.51$

6) Scenario 8 (SATS, approach to missed approach with virtual traffic), $M = 7.38$, $SD = 1.98$

To determine if evaluation pilots reported different SART ratings when different types of scenarios were performed using different types of procedures, a series of Wilcoxon tests was performed. Because the differences between the SART ratings associated with particular pairs of scenarios were of primary interest (i.e., scenarios 3 vs. 4, 5 vs. 6, and 7 vs. 8), only the results of the Wilcoxon tests associated with these scenario pairings are discussed here.

As shown by the mean SART scores plotted in Fig. 12, evaluation pilots reported higher SART scores when scenarios were performed using the SATS procedures than when scenarios were performed using the baseline procedures.

The Wilcoxon tests revealed that, at a statistically significant level, evaluation pilots reported higher SART scores when performing scenario 6 as compared with scenario 5 ($p = 0.0161$). Statistically, evaluation pilots reported equivalent SART ratings when performing scenario 3 as compared with scenario 4 ($p = 0.2513$) and when performing scenario 7 as compared with scenario 8 ($p = 0.3969$). In

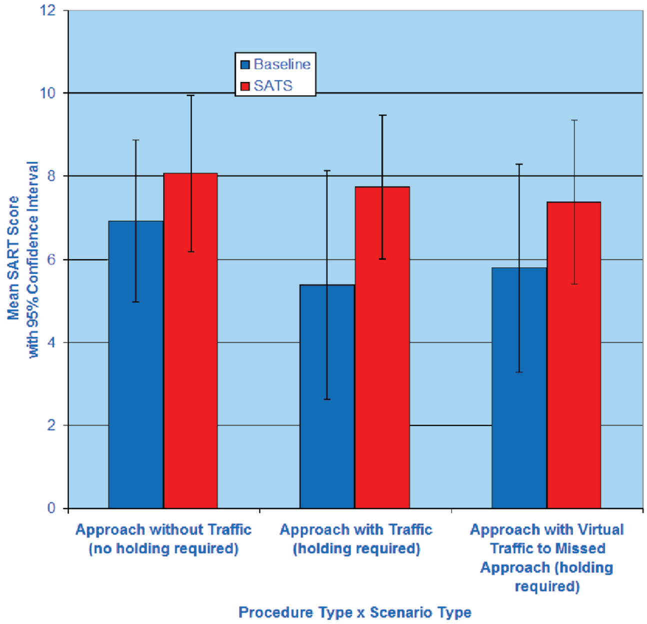

Fig. 12 Mean SART ratings associated with procedure type × scenario type.

summary, the situation awareness hypothesis that "equivalent situation awareness ratings will be associated with all scenarios" was supported.

D. Subjective Assessment of Traffic Awareness

Evaluation pilots reported the following traffic awareness ratings when they performed different types of scenarios using baseline and SATS procedures ($N = 12$ for each scenario):

1) Scenario 3 (baseline, approach without traffic), $M = 5.25$, $SD = 1.59$

2) Scenario 4 (SATS, approach without traffic), $M = 6.13$, $SD = 1.23$

3) Scenario 5 (baseline, approach with virtual traffic), $M = 5.54$, $SD = 1.50$

4) Scenario 6 (SATS, approach with virtual traffic), $M = 6.64$, $SD = 0.40$

5) Scenario 7 (baseline, approach to missed approach with virtual traffic), $M = 5.54$, $SD = 1.42$

6) Scenario 8 (SATS, approach to missed approach with virtual traffic), $M = 6.42$, $SD = 0.42$

To determine if evaluation pilots reported different traffic awareness ratings when different types of scenarios were performed using different types of procedures, a series of Wilcoxon tests was performed. Because the differences between the traffic awareness ratings associated with particular pairs of scenarios were of primary interest (i.e., scenarios 3 vs. 4, 5 vs. 6, and 7 vs. 8), only the results of the Wilcoxon tests associated with these scenario pairings are discussed here.

As shown by the mean traffic awareness scores plotted in Fig. 13, evaluation pilots reported higher traffic awareness scores when scenarios were performed using the SATS procedures than when scenarios were performed using the baseline procedures.

The Wilcoxon tests revealed that, at a statistically significant level, evaluation pilots reported higher traffic awareness scores when performing scenario 6 as compared with scenario 5 ($p = 0.0418$) and when performing scenario 8 as compared with scenario 7 ($p = 0.0117$). Statistically, evaluation pilots reported equivalent traffic awareness ratings when performing scenario 3 as compared with scenario 4 ($p = 0.0581$). In summary, the situation awareness hypothesis that "equivalent situation awareness ratings will be associated with all scenarios" was supported.

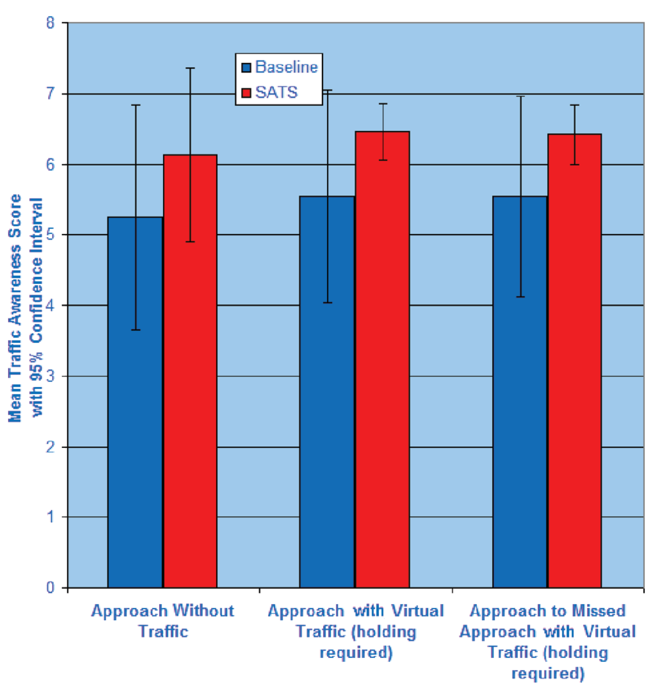

Fig. 13 Mean traffic awareness ratings associated with procedure type × scenario type.

E. Subjective Assessment of Navigation Guidance Awareness

Evaluation pilots reported the following ratings for navigation guidance SA when they performed different types of scenarios using baseline and SATS procedures ($N = 12$ for each scenario):

1) Scenario 3 (baseline, approach without traffic), $M = 5.63$, $SD = 1.17$

2) Scenario 4 (SATS, approach without traffic), $M = 6.50$, $SD = 0.43$

3) Scenario 5 (baseline, approach with virtual traffic), $M = 5.04$, $SD = 1.63$

4) Scenario 6 (SATS, approach with virtual traffic), $M = 6.50$, $SD = 0.43$

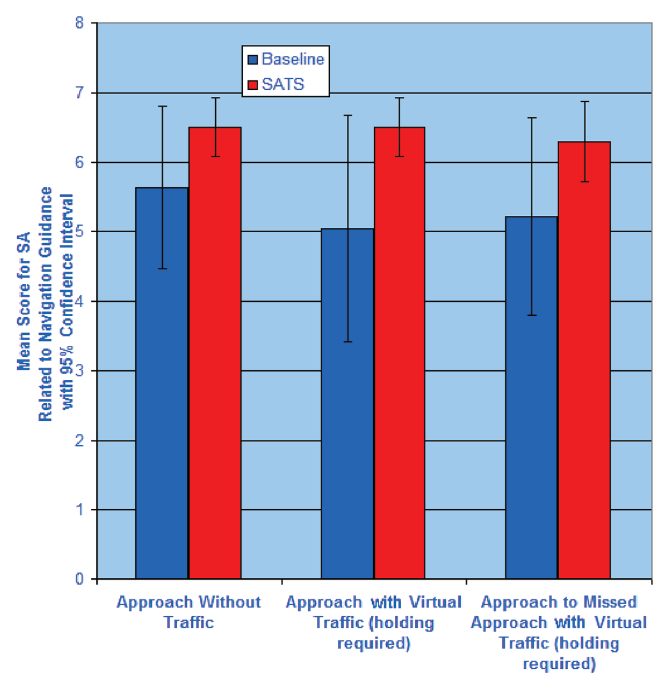

Fig. 14 Mean ratings for SA related to navigation guidance associated with procedure type × scenario type.

5) Scenario 7 (baseline, approach to missed approach with virtual traffic), $M = 5.21$, $SD = 1.42$

6) Scenario 8 (SATS, approach to missed approach with virtual traffic), $M = 6.29$, $SD = 0.58$

To determine if evaluation pilots reported different ratings for navigation guidance SA when different types of scenarios were performed using different types of procedures, a series of Wilcoxon tests was performed. Because the differences between the ratings for navigation guidance SA associated with particular pairs of scenarios were of primary interest (i.e., scenarios 3 vs. 4, 5 vs. 6, and 7 vs. 8), only the results of the Wilcoxon tests associated with these scenario pairings are discussed here.

As shown by the mean scores plotted in Fig. 14, evaluation pilots reported higher ratings for navigation guidance SA when scenarios were performed using the SATS procedures than when scenarios were performed using the baseline procedures.

The Wilcoxon tests revealed that, at a statistically significant level, evaluation pilots reported higher ratings for navigation guidance SA when performing scenario 4 as compared with scenario 3 ($p = 0.0486$) and when performing scenario 6 as compared with scenario 5 ($p = 0.0098$). Statistically, evaluation pilots reported equivalent ratings for navigation guidance SA when performing scenario 7 as compared with scenario 8 ($p = 0.0550$). In summary, these data illustrate that pilots reported their navigation guidance awareness increased when they flew SATS HVO scenarios (surpassing the situation awareness hypothesis that "equivalent situation awareness ratings will be associated with all scenarios").

V. Conclusion

Detailed results of pilots' flight technical error and their subjective assessments of workload and situation awareness indicate that the SATS HVO concept is a viable improvement to the procedural control operations in use today at nontowered, nonradar-controlled airfields in instrument meteorological conditions. The HVO Flight Experiment, flown on NASA's Cirrus SR22, used a subset of the HVO Simulation Experiment scenarios and evaluation pilots in order to validate the simulation experiment results. The evaluation pilots easily transitioned their simulator experience and familiarity with HVO procedures to the flight deck.

Results of the subjective assessments revealed that all 12 low-time instrument-rated pilots preferred SATS HVO when compared with current procedural separation operations. Pilots expressed their frustration with the lengthy hold maneuvers on the baseline scenarios and their relief at being able to fly the more efficient SATS approaches. Evaluation pilots maintained airspeed and lateral path more accurately when they performed the SATS scenarios than when they performed the baseline scenarios. They also maintained altitude equally well in both SATS and baseline scenarios. By observations of lateral path error data, the significant pilot improvement for the SATS scenarios is intriguing, because the flight guidance system (HSI and multifunction display with moving map and traffic) was identical for both baseline and SATS scenarios. Although not the focus of this study, it can be surmised that the pilots were more "engaged" in the SATS arrival sequence process, anticipating their next maneuver, instead of reacting to ATC clearance, and thus they flew more precisely. The notion that the SATS HVO flight procedure itself produces better flight performance was a surprising observation and merits further study. Evaluation pilots also assessed their workload to be lower in SATS scenarios, and increased situation awareness with respect to traffic and navigation guidance.

Supporting the SATS pilot has been a goal of the research and although minimally qualified (low-time instrument-rated) pilots were the evaluation pilots in the HVO Flight Experiment, more experienced and qualified pilots should be able to fly SATS HVO as well. The HVO flight tasks were reviewed by evaluation pilots as a logical extension of the instrument rating so should easily merge with FAA training and certification curriculum without adding significant requirements. Evaluation pilots also performed HVO tasks well with the simple NASA Cirrus SR22 research avionics configuration, and so improved avionics (including a primary flight display) may simplify tasks further.

The piloted experiment described in this paper was a critical part of the building-block validation and verification process of SATS HVO that included multiple elements ranging from concept development to full-system architecture proof-of-concept demonstration that was successfully shown to the public at the June 2005 SATS Technical Demonstration in Danville, Virginia.

Acknowledgements

The authors wish to thank and recognize those who have contributed to the SATS HVO concept evaluation and the development of this paper. Specifically we would like to thank Maria Consiglio, Brian Baxley, Terry Abbott, Ken Jones, Gary Lohr, Victor Carreno, Cesar Munoz, Sheila Conway, Sally Johnson, Hazari Syed, and Jim Sturdy for their insight and advice.

References

[1] Long, D., Lee, D., Johnson, J., and Kostiuk, P., "A Small Aircraft Transportation System (SATS) Demand Model," NASA CR-2001-210874, June 2001.

[2] Free Flight Steering Committee, "National Airspace System; Concept of Operation and Vision for the Future of Aviation," Radio Technical Commission for Aeronautics, Washington, DC, Nov. 2002.

[3] Abbott, T., Jones, K., Consiglio, M., Williams, D., and Adams, C., "Small Aircraft Transportation System, Higher Volume Operations Concept: Normal Operations," NASA TM-2004-213022, May 2004.

[4] Dowek, G., Munoz, C., and Carreno, V., "Abstract Model of the SATS Concept of Operations: Initial Results and Recommendations," NASA TM-2004-213006, 2004.

[5] Williams, D. M., Consiglio, M. C., Murdoch, J. L., and Adams, C. H., "Preliminary Validation of the Small Aircraft Transportation Systems Higher Volume Operations (SATS HVO) Concept," *Proceedings of the 24th International Congress of the Aeronautical Sciences*, National Technical Information Service, Springfield, VA, 2004, pp. 1–10.

[6] Consiglio, M., Williams, D., Murdoch, J., and Adams, C., "SATS HVO Normal Operations Concept Validation Simulation Experiment," *AIAA/ATIO Conference*, AIAA, Reston, VA, 2005

[7] Williams, D., Consiglio, M., Murdoch, J., and Adams, C., "Flight Technical Error Analysis of the Small Aircraft Transportation System Higher Volume Operations (SATS HVO) Simulation and Flight Experiments," *24th DASC: the 24th Digital Avionics Systems Conference*, Vol. 2, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 2005, pp. 1–12.

[8] Consiglio, M., Conway, S., Adams, C., and Syed, H., "SATS HVO Procedures for Priority Landings and Mixed VFR/IFR Operations at Non-Towered, Non-Radar Airports," *24th DASC: the 24th Digital Avionics Systems Conference*, Vol. 2, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 2005, pp. 1–8.

[9] Adams, C. A., Consiglio, M. C., Conway, S., and Syed, H., "The Pilot Advisor: Assessing the Need for a Procedural Advisory Tool," *24th DASC: the 24th Digital Avionics Systems Conference*, Vol. 2, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 2005, pp. 1–8.

[10] Magyarits, S., Racine, N., and Hadley, J., "Air Traffic Control Feasibility Assessment of Small Aircraft Transportation System (SATS) Higher Volume Operations (HVO)," DOR/FAA/CT-05/26, National Technical Information Service, Springfield, VA, May 2005.

[11] Anon., "Instrument Rating Practical Test Standards for Airplane, Helicopter, Powered Lift," Federal Aviation Administration, FAA-S-8081-4D, 2004.

[12] Norusis, M. J., SPSS® for Windows™: Base System User's Guide, Ver. 6.0, SPSS, Chicago, 1993.

[13] Keppel, G., *Design and Analysis: A Researcher's Handbook*, 3rd ed., Prentice–Hall, Upper Saddle River, NJ, 1991.

[14] Wierwille, W. W., and Casali, J. G., "A Valid Rating Scale for Global Mental Workload Measurement," *Proceedings of the Human Factors Society 27th Annual Meeting*, Human Factors Society, Santa Monica, CA, 1983, pp. 129–133.

[15] Wilcoxon, F., "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, Vol. 1, 1945, pp. 80–83.

[16] Taylor, R. M., "Situational Awareness Rating Technique (Sart): The Development of a Tool for Aircrew Systems Design," *In Situational Awareness in Aerospace Operations*, AGARDograph CP-478, NATO-AGARD, Euilly Sur Seine, France, 1990, pp. 3/1–3/17.