

Understanding Crew Decision Making in the Presence of Complexity – A Flight Simulation Experiment

Steven D. Young¹, Taumi S. Daniels², and Emory (Tom) Evans³
NASA Langley Research Center, Hampton, VA, 23681

Maarten Uijt de Haag⁴ and Pengfei (Phil) Duan⁵
Ohio University, Athens, OH, 45701

Crew decision making and response have long been leading causal and contributing factors associated with aircraft accidents. Further, it is anticipated that future aircraft and operational environments will increase exposure to risks related to these factors if proactive steps are not taken to account for ever-increasing complexity. A flight simulation study was designed to collect data to help in understanding how complexity can, or may, be manifest. More specifically, an experimental apparatus was constructed that allowed for manipulation of information complexity and uncertainty, while also manipulating operational complexity and uncertainty. Through these manipulations, and the aid of experienced airline pilots, several issues have been discovered, related most prominently to the influence of information content, quality, and management. Flight crews were immersed in an environment that included new operational complexities suggested for the future air transportation system as well as new technological complexities (e.g. electronic flight bags, expanded data link services, synthetic and enhanced vision systems, and interval management automation). In addition, a set of off-nominal situations were emulated. These included, for example, adverse weather conditions, traffic deviations, equipment failures, poor data quality, communication errors, and unexpected clearances, or changes to flight plans. Each situation was based on one or more reference events from past accidents or incidents, or on a similar case that had been used in previous developmental tests or studies. Over the course of the study, 10 two-pilot airline crews participated, completing over 230 flights. Each flight consisted of an approach beginning at 10,000 ft. Based on the recorded data and pilot and research observations, preliminary results are presented regarding decision-making issues in the presence of the operational and technological complexities encountered during the flights.

Nomenclature

<i>ADS-B</i>	=	Automatic Dependent Surveillance – Broadcast
<i>AIS/MET</i>	=	Aeronautical and Meteorological Information Service
<i>A-SMGCS</i>	=	Advanced Surface Movement Guidance and Control Systems
<i>ASTAR</i>	=	Airborne Spacing for Terminal Arrival Routes
<i>ATC</i>	=	Air Traffic Control
<i>ATIS</i>	=	Automatic Terminal Information Service
<i>CMF/RFD</i>	=	Cockpit Motion Facility's Research Flight Deck simulator
<i>DataComm</i>	=	Data Communications (between pilots and air traffic control)
<i>EFB</i>	=	Electronic Flight Bag
<i>EFVS</i>	=	Enhanced Flight Vision Systems
<i>EVO</i>	=	Equivalent Visual Operations
<i>FAF</i>	=	Final Approach Fix

¹ Research Engineer, Safety-Critical Avionics Systems, 1 S. Wright St., Associate Fellow.

² Research Engineer, Electromagnetics and Sensors, 8 N. Dryden St., Member.

³ Research Engineer, Flight Software Systems, 5 N. Dryden St., Member.

⁴ Edmund K. Cheng Professor, School of Electrical Engineering, Stocker Center, Associate Fellow.

⁵ Graduate Student, School of Electrical Engineering, Stocker Center, Student Member.

<i>FIM</i>	= Flight Deck Interval Management
<i>HUD</i>	= Head-Up Display
<i>IB</i>	= Inattential Blindness
<i>ILS</i>	= Instrument Landing System
<i>KMEM</i>	= Memphis International Airport
<i>LNAV</i>	= Lateral Navigation
<i>MSL</i>	= Mean Sea Level
<i>NACp</i>	= Navigation Accuracy Category for position
<i>NACv</i>	= Navigation Accuracy Category for velocity
<i>NCO</i>	= Net-Centric Operations
<i>ND</i>	= Navigation Display
<i>NextGen</i>	= The U. S. Next Generation Air Transportation System
<i>NOTAM</i>	= Notice to Airmen
<i>PFD</i>	= Primary Flight Display
<i>RNAV</i>	= Area Navigation
<i>TCAS</i>	= Traffic Alerting and Collision Avoidance System
<i>VNAV</i>	= Vertical Navigation

I. Introduction and Motivation

Understanding and improving crew decision making and response in the presence of complexity is a technical challenge being investigated by the Vehicle Systems Safety Technology project within NASA's Aviation Safety Research Program. This technical challenge is motivated in large part by leading causal and contributing factors associated with aircraft accidents. For example, ineffective alerting mechanisms can lead to less than optimal pilot decision making and result in aircraft loss of control accidents¹. A second, and equally important motivation for research addressing this challenge is a recognition that future aircraft and operational environments will likely increase exposure to risk related to these factors if proactive steps are not taken to account for ever-increasing complexity. In the U.S., NextGen² is envisioned to provide a safe, efficient and reliable air transportation system for the future that removes many of the constraints in the current system, supports a wider range of operations, and thus delivers an overall system capacity well above that of current operating levels. However, few would dispute that with these new capabilities comes additional complexity.

The research foci and scope of the NASA project's activities addressing this technical challenge are two-fold. First, the project seeks to proactively identify and mitigate NextGen flight deck risks, particularly those associated with piloting during trajectory-based operations in terminal areas in reduced visibility or at night. Second, the project seeks to develop and evaluate interventions that address documented crew-related causal factors of accidents, with specific attention paid to pilot proficiency in increasingly automated flight decks (e.g. manual handling skill erosion), spatial disorientation, and loss of energy state awareness leading to loss of control.

This paper describes one activity under this broad scope – results of a flight simulation study designed to collect data to help in understanding how complexity affects crew decision making. Through manipulations of complexity, and the aid of experienced airline pilots, decision-making issues are discovered, and the influence of information content, quality, and management are observed. As a secondary objective, the usability of two new display concepts is assessed for their ability to improve decision-making effectiveness. Only the discovery aspects of the study are presented here. The usability assessments will be reported elsewhere.

The term “complex” can be defined to have many properties; two are used here: (1) difficult to understand; and (2) made up of many interrelated parts. Complexity is considered an overarching trait associated with a flight crew's work domain and can result from either, or both, of these properties. Compounding the problem is the fact that there are multiple facets to the pilot's work domain that may be “difficult to understand” or involve “many interrelated parts”. Four facets are addressed in this study: complexity in operations, complexity in systems and procedures, complexity in the flight environment, and complexity due to uncertainty. A fifth facet is not addressed: complexity in training.

To explore this complexity space with the goal of better understanding decision-making issues, a set of commercial airline flight crews were immersed in a NextGen-like environment – one of the most complex operational and technological environments ever conceived. Several new concepts were included that are under development in the aviation industry, but are at various stages of technical maturity and insertion into operations. This enabled the study to address the first two facets. In addition, a set of off-nominal situations were emulated (e.g. weather conditions and system failures). This enabled the study to address the second two facets.

A. Complexity in Operations

Complexity in operations was created based upon current commercial airline operations, then adding aspects of several concepts under development. A terminal area operational environment was created in a flight simulator for the testing based upon previous work³, and using the Memphis International Airport (KMEM). This environment included aspects of: closely-spaced parallel approaches, flight deck interval management (FIM), advanced surface movement guidance and control systems (A-SMGCS), pilot-controller data communications (DataComm), equivalent visual operations (EVO), and net-centric operations (NCO). Each of these requires changes to roles, responsibilities, level of automation support, and procedures for both air traffic controllers and flight crews.

Closely-spaced parallel approach operations were emulated at KMEM in four forms depending on the scenario chosen: (1) simultaneous landings to runways 36L/36C, with 3 nmi in-trail spacing, a 1.5 nmi stagger, and no departures; (2) simultaneous landings to runways 36L/36C, with 5 nmi in-trail spacing, a 2.5 nmi stagger, and no departures; (3) the same as (1) but using runways 18C/18R; and (4) the same as (2), but with departures from runway 36R. Also, in terms of ATC, operations to the two parallel runways were controlled independently (e.g. voice/data communications on separate frequencies). Runways 36L/18R and 36C/18C are separated by 3500 feet.

FIM is within the “Time-Based Flow Management” implementation portfolio of Ref. 2 and is intended to help assure a smooth flow of traffic and an increase in efficiency. FIM operations were emulated in this study based upon the current standards⁴, and as extended by previous research^{3,5,6}. FIM is enabled using transponder data from another aircraft (e.g. aircraft position and identification) and an onboard system function that provides speed suggestions, or guidance, to the crew or autopilot function. When these speeds are followed, an in-trail spacing goal will be met at a pre-selected “achieve by” point. In this study, this point was set as the final approach fix (FAF). The aircraft to follow is assigned by Air Traffic Control (ATC) and called the “target.” This target may be on a different arrival path early in the flight, but after passing a merge point (usually near the outer marker), the two aircraft are paired in-trail. The speed profile is continuously calculated throughout the flight, and may differ from charted speeds at waypoints along the path. In this study the computed speed suggestions were given on the Primary Flight Display (PFD) and the pilot’s task was to control the aircraft’s speed to match these suggestions, as long as the speeds were deemed safe and appropriate (e.g. would not lead to a loss of separation). Status with respect to achieving the spacing goal at the FAF was given on the Navigation Display (ND). The in-trail spacing goals for this study were either (1) 78 seconds, or (2) 150 seconds at the FAF. These temporal spacing goals corresponded to the 3 and 5 nmi spacing cases, respectively. See the Appendix for the FIM display symbology provided to the crews in this study.

The A-SMGCS concept provides the necessary routing, guidance and surveillance for the control of aircraft and vehicles in order to maintain a desired surface movement rate under all weather conditions⁷. In Ref. 2, this is part of the “Improved Surface Operations” implementation portfolio. Aspects of A-SMGCS operations were emulated based on previous research^{3,8,9} and published requirements¹⁰, and included: (1) an airport moving map display, available on both the ND and the Electronic Flight Bag (EFB); (2) synthetic vision-based airport features and surface guidance on the PFD, (3) runway incursion detection functions, providing indications and alerts on all displays as well as in audio form, and (4) DataComm, as described below. See the Appendix for the display symbology provided to the crews in this study in support of A-SMGCS.

DataComm is an operating concept to supplement existing voice communications and provide two-way data exchange between ATC and flight crews for clearances, instructions, advisories, flight crew requests, and reports¹¹. DataComm is required by multiple implementation portfolios in Ref. 2. DataComm operations were emulated in this study by creating a mixed voice/data communications environment in the simulator. Voice instructions were limited to frequency changes and clearances (e.g. “cleared to land”), and these instructions were given to all traffic. Voice was also used for any off-nominal requests from the pilots (e.g. requesting to avoid weather, or to go-around). This is also known as the “voice by exception” concept. Data communications were used for all required interchanges, with unique messages required to support FIM and A-SMGCS. All messages from ATC were displayed on the PFD and on the EFBs. The EFBs were used for reviewing and accepting/rejecting the ATC-provided messages. Each incoming message was announced via a chime in the flight deck. See the Appendix for the DataComm display symbology provided to the crews in this study, and examples of transmitted messages.

EVO alludes to the ability to maintain operational tempos at airports regardless of weather/visibility conditions. In Ref. 2, this goal is addressed within the “Improved Approaches and Low-Visibility Operations” implementation portfolio, which depends on advances in Enhanced Flight Vision Systems (EFVS) as an enabling technology. In this study, an EFVS was implemented based on previous research and consistent with existing standards¹². Through use of the EFVS, the crews were permitted to continue approaches in low visibility conditions and with the reduced separations described above. Enabling autoland was not required; although some crews decided to enable it for some flights. See the Appendix for examples of the EFVS display symbology provided to the crews in this study.

NCO encourages a greater sharing of information across the many agents within the airspace system, with each aircraft comprising one such agent. This goal is enabled most prominently through the use of data link(s) above and beyond just pilot-ATC DataComm (as described above). In Ref. 2, this capability is being developed within the “On Demand National Airspace System Information” implementation portfolio. For this study, aspects of NCO were emulated by the Aeronautical and Meteorological Information Service (AIS/MET) data link and associated displays and interaction via the EFB. This data link is envisioned as an evolution of the current Aircraft Communications Addressing and Reporting System and Flight Information System Broadcast data links, and reacts in part to Ref. 13 by “improving the system of providing airmen with pertinent and timely information regarding the national airspace system” and by “decreasing the overwhelming volume of Notices to Airmen (NOTAMs) an airman receives when retrieving information prior to a flight.” The implementation in this study followed Ref. 14 and emulated several of the use cases specified there for the approach and landing phases of flight. This includes the transmission of significant NOTAMs and/or weather information relevant to the flight.

B. Complexity in Systems and Procedures

Complexity in systems and procedures was created based upon requirements for the operations described above, and emerging technologies which are becoming prevalent. For this study, standard avionics systems and autopilot functions were employed. Additional technologies were integrated, or modeled. These included: the Airborne Spacing for Terminal Arrival Routes (ASTAR) function, Automatic Dependent Surveillance – Broadcast (ADS-B) transponders and associated displays of traffic information, EFBs, EFVS, AIS/MET data link, airport moving map functions and displays, and runway incursion indication and alerting functions. A brief description of each of these systems is provided below. See the Appendix for the associated display symbology that was provided to the crews.

ASTAR^{5,6,15} is an onboard algorithm (i.e. automation) that supports the FIM operation by constantly calculating the airspeed required to position the aircraft at an “achieve by” point in accordance with a desired spacing goal behind a target aircraft. Several years of development have culminated in the version of ASTAR used for this experiment (v11.0). For this study, the ASTAR function provided the crew with speed guidance on the PFD to optimize the aircraft’s velocity profile during the approach in order to achieve an in-trail spacing goal at the FAF. In addition, ASTAR provided the crew with status information regarding in-trail spacing on both the ND and the Control Display Unit. Symbology was based on Ref. 3 with only minor modifications.

ADS-B is a cooperative surveillance system that requires aircraft transponders capable of transmitting global positioning system-based velocity and position information at prescribed rates. Such transponders have recently been mandated by the FAA (a.k.a. “ADS-B out”). For this experiment, the onboard system was also able to receive ADS-B messages from other aircraft and pass them to display functions in order to depict the location of relevant traffic. This “ADS-B In” capability is planned as part of the evolution of the air transportation system.

EFBs are used in current airline operations, but only on a relatively small scale, and with somewhat limited functionality with respect to their potential. For this experiment, a baseline set of functions was implemented to be consistent with current state-of-the-practice. Functionality was added to support the above-mentioned operating concepts. From the top-level (Main Menu) page of the EFB, seven categories of information were available via sub-menus. These were: Communications, Weather, NOTAMs, Charts, Airport Moving Map, Checklists, and Automatic Terminal Information Service (ATIS). Two EFBs were installed, left and right, to allow either pilot to access information independently. Procedures for use of the EFBs were not prescribed, with only a few exceptions: (1) on receipt of a data link message, either from ATC or via the AIS/MET service, one of the pilots was required to review the message, discuss it with the other crewmember, then reply using the EFB (e.g. via a “wilco” message); (2) prior to commencement of each flight, the crew was required to review and discuss current NOTAMs and weather information stored in the EFB; and (3) the crew was required to utilize the charts and checklists in the EFB per their typical airline procedures. This last task was in support of the “paperless cockpit” concept¹⁶. No paper reference material was available during the experiment, other than a few pages from the operating manual and quick reference handbook that were occasionally needed to support emergency checklist items.

EFVS (a.k.a. synthetic and enhanced vision systems) technologies are certified on aircraft, but are not widely used to date. However, it is presumed that some form of these systems will be required to enable low visibility operations without reducing traffic flow rates. For this experiment, the EFVS was based on Ref. 3 and provided the crews with an ego-centric view of significant airport features while on final approach and during taxi. These features included: runways, taxiways, centerlines, edge lines, markings, and vertical structures (e.g. buildings). On the PFD, this was presented using synthetic vision. On the HUD, this was based upon enhanced vision technology (i.e. an infrared sensed image). There were no changes to procedures due use of the EFVS.

For this experiment, the AIS/MET data link technology was implemented to comply with the newest standards¹⁴, with the pilot interface via the EFB and ND. The type(s) of AIS/MET messages sent during a flight over the course of the experiment were drawn from the following:

- Temporary obstruction (text and graphic)
- Operationally-significant NOTAMs (text and graphic)
- Changes to instrument approach procedures (text and graphic)
- Runway braking action reports (text only, no graphic)
- Runway winds (text only, no graphic)
- Pilot reports (text only, no graphic)

The airport moving map functions and displays were implemented as in Ref. 3, with some changes and additional features. As shown in the Appendix, the moving map depicted locations of runway and taxiway edges, painted markings, and buildings, with overlays of aircraft position, traffic positions, and the ATC-provided taxi route. This latter was provided via data communications. The moving map was integrated with the NOTAM system such that any geo-referenced NOTAMs were displayed graphically (e.g. closed runways). Runway status was indicated as described below. The moving map function was available to the crew on both the ND and the EFB.

Runway incursion detection and alerting functions were implemented in accordance with Ref. 10. These functions provide a level of protection by warning pilots if the runway is occupied by other traffic. More specifically, for this experiment, while on final approach, if traffic was ahead of the aircraft and on the runway, the following indications were provided to the crew:

- When 2 nmi from the runway threshold
 - Textual “RSI” shown on the ND
 - Blue runway outline on PFD and ND
- When within 35 sec of the runway threshold
 - Audible warning: “CAUTION TRAFFIC”
 - Yellow runway outline on PFD and ND
- When within 15 sec of the runway threshold
 - Audible warning: “WARNING TRAFFIC”
 - Red runway outline on PFD and ND

In addition, after landing, if the same conditions above were true for the aircraft that was landing next, then the same indications were provided to the aircraft rolling out until exiting the runway.

Area Navigation (RNAV) arrival procedures were used for all flights. These were based on the published procedures for the KMEM airport. Scenarios were defined to begin on one of four RNAV-based paths, each of which transitioned to use Instrument Landing System (ILS) guidance near the outer marker. At the beginning of each run (10,000 ft MSL), pilots engaged Lateral Navigation (LNAV) and Vertical Navigation (VNAV) modes in the autopilot such that the aircraft followed a path corresponding to the published RNAV arrival procedure. The four nominal paths enabled flights to approach KMEM from the four ordinal directions (Northwest, Northeast, Southwest, and Southeast).

In terms of general procedures, the crews were instructed to treat each flight as a commercial flight, realizing that the simulator and the environment requires some adjustments to what is normally done (e.g. reviewing NOTAMs on the EFBs). In addition, since pilots were paired (Captain, First Officer) according to airline, crews were instructed to use their company’s best practices for crew resource management. Finally, time was allocated before flying for the crew to discuss and agree on responsibilities, and then to perform practice flights as well. Pilot-flying and pilot-monitoring roles were occasionally switched; however, pilots did not change seats.

C. Complexity in the Flight Environment

For this study, flight environment complexity encompassed various weather conditions and aircraft conditions. Three weather conditions were emulated. For the first, the visibility was set to 3 nmi and the cloud ceiling was set to 1000 ft. Time of day was set to early evening (dusk), winds were light, and there was no turbulence or rain. For the second and third conditions, storm models were used based on data provided by NOAA for two events that occurred at KMEM in 2011, and for which the airport had significant weather-induced delays. Each model included 3D winds, radar reflectivity, and icing estimates. The nature of the two storms was quite different in terms of intensity and variability, providing a good span of conditions to test within. These models were used to drive the motion-platform (e.g. winds and turbulence effects), and to determine what information should be provided to the crews (e.g. on the EFBs and weather radar display) during the flight.

The off-nominal aircraft conditions that were tested included: (1) a failure of the anti-skid system, (2) a loss of hydraulic pressure, and (3) a loss of the ILS signals. Each of these are situations that airline crews regularly train for, and have well-defined procedures to follow should they occur. These procedures (checklists) were available to the crew on request during the flights.

D. Complexity Due to Uncertainty

Traffic deviations, poor data quality, and communication errors were used in the study to manipulate complexity due to uncertainties. While nominal flight tracks were followed by terminal area traffic, aircraft would occasionally deviate from the nominal track per unique scenario designs. Examples include an aircraft heading to the wrong runway, an aircraft reaching waypoints ahead or behind schedule, or an aircraft unexpectedly slowing either in flight or during landing roll-out. Different levels of errors in the ADS-B position and velocity reports, as well as ownship position and velocity estimates, were also modeled. ADS-B errors were based on the performance parameters: Navigation Accuracy Category for position (NACp) and Navigation Accuracy Category for velocity (NACv). Ownship position and velocity errors were based on typical error characteristics for the Global Navigation Satellite Systems and Inertial Navigation Systems. Depending on the magnitude of modeled errors, airborne conflict and runway incursion detection algorithms will have higher/lower false alarm and missed detection rates.

Poor data quality was also reflected in the weather information products. Because of the movement of the weather events, model uncertainty, weather sensors/radar performance, and data link performance, inconsistencies between the information provided by the weather information displays were introduced. For example, the weather radar observations during the flights were translated and altitude-variant with respect to the ground-based radar images available on the EFB. This to account for the movement of the weather and the aircraft.

Communications errors were introduced in both the voice and data modes for selected flights, including discrepancies between ATC voice and data instructions.

Specifics as to which of the above uncertainties would occur, and when, were not briefed; however, crews were briefed that such situations may occur during the experiment.

E. Manifestations of Complexity

The following unsafe situations can manifest themselves in the presence of flight deck complexity during terminal area operations: runway incursions and excursions, spatial disorientation, unstable approaches or loss of energy state awareness, automation mode confusion, wake encounters, path deviations, and loss of separation from traffic, terrain, or airspace constraints. All of these situations have been associated with past accidents and incidents, and many of the above-mentioned operational concepts and technologies are aimed at reducing their likelihood, or mitigating their effects. Given the totality of the complexities described in the previous subsections, and the requisite training and transition processes, it is unknown whether the risk of these unsafe situations manifesting themselves will reduce, increase, or stay the same. It is also unknown whether new manifestations will emerge. Research such as the experiment conducted here should help to uncover some of these unknowns.

II. Experiment Design

The experiment was decomposed into three parts, each with a dedicated block of time in the simulator. Further each flight crew was available for two full working days. The nominal schedule for each crew is shown in Table 1.

Part 1 of the experiment was a usability assessment of the EFBs for providing crews with textual and graphic NOTAMs, weather, and information provided by the AIS/MET data link. The results of this assessment will be reported elsewhere; however, because these same EFBs and EFB functions were also used for Part 2, some explanation is necessary, and is given in the Appendix. Note that some information was provided as a pre-flight load (e.g. to support pre-flight briefings of weather and NOTAMs), and other information was provided in-flight via data link per the NCO concept. Part 3 was a usability assessment of additional display concepts for the PFD/ND. Because the Part 3 display concepts were not used during Part 2, they are not discussed here.

Part 2 focused on the discovery objective described in Section I – to collect data to help in understanding how complexity affects crew decision making. In this Part, complexity is manipulated and experienced airline pilots help to discover decision-making issues, and to assess the influence of information content, quality, and management. Findings here serve to identify areas requiring further research, or a re-definition of designs or operating concepts.

With the exception of the EFB (see Appendix), flight deck display functions were based on previous work, starting with the capability developed for Ref. 3, and extending as necessary to remain consistent with industry developments and standards. To provide the required information to the display functions, a flexible comprehensive avionics system model¹⁷ was integrated and adapted. This model provided hazard and integrity monitoring functions

as well as integrated alerting and notification capabilities. This model included sub-models for avionics technologies such as ADS-B and weather radar, and allowed for testing across a number of experimental conditions regarding information system performance and data quality. Also, a high-fidelity atmospheric environment was realized by using numerical weather data sets derived from NOAA’s High-Resolution Rapid Refresh model. These data sets coincided with severe weather events that occurred at the test airport (KMEM) in 2011. The NOAA model provided estimates in three dimensions for radar reflectivity, icing, turbulence, and winds. These data were used to affect the aircraft aerodynamic model, the onboard weather radar system, and the weather information provided via data link¹⁸.

Table 1. Crew Schedule

800	Arrive facility	800	Arrive facility
815	Eye tracker calibration	820	Part 2 (4-6 approaches)
830	Welcome & overview	1000	Break
900	Training (Classroom)	1010	Part 2 (2-4 approaches)
1000	Break	1100	Part 2 Debrief
1015	Sim. Familiarization	1200	Lunch
1115	Lunch	1245	Part 3 (training & practice)
1215	Part 1 (3-6 approaches)	1315	Part 3 (2-4 approaches)
1330	Part 1 Debrief	1415	Break
1400	Part 2 (4-6 approaches)	1430	Part 3 (2-4 approaches)
1530	Break	1620	Part 3 Debrief
1540	Part 2 (2-4 approaches)	1700	End of Day 2
1700	End of Day 1		

A. Experiment Apparatus

The experiment was hosted by the Cockpit Motion Facility’s Research Flight Deck (CMF/RFD) at NASA’s Langley Research Center. The physical structure is shown in Fig. 1. This platform has been designed to test state-of-the-art technologies while providing a baseline similar to current commercial transport aircraft. The simulator has many commonalities with the B-757/767 including: the aerodynamic model and handling qualities, the flight management system, the control display unit, the automation mode control panel, the center aisle stand and throttle quadrant, and the overhead panel. More consistent with the B-787 are the four 17-inch liquid crystal displays, the dual HUDs, and the EFBs. Dual sidesticks provide similarity with Airbus aircraft.

An additional instrumentation system was utilized to perform head and eye tracking. This system (Fig. 2) recorded data on the look angle and dwell times for both crewmembers during the experiment.

Simulation control stations were also utilized to: (1) setup and monitor each flight, (2) host air traffic controllers and control functions, and (3) host the avionics system model¹⁷ and weather environment models¹⁸.

B. Experiment Variables and Data Recording

The independent variables for the experiment are enumerated in Table 2. Data was collected in many forms including interview and questionnaire responses, audio/video recordings, aircraft state and performance data files, pilot input data files (e.g. button presses and flight control inputs), head and eye tracking data files, and pilot/researcher observations. Short questionnaires were completed by each pilot following each flight. The summary debriefing and interview following Part 2 lasted about 1 hour.

C. Test Conditions and Scenarios

Each crew completed a set of flights into the KMEM airport. All flights began at 10,000 ft MSL, on one of the four RNAV-based arrival and approach paths, and at an airspeed of 250 knots. Flights ended nominally after landing and runway exit; however, there were many cases where the crew decided to go-around. The level and type of complexity were manipulated by defining 22 specific scenarios based upon setting different values for the independent variables discussed previously and listed in Table 2. Scenarios were designed and selected such that crews would be exposed to a span of situations or conditions, each based on one or more reference events from past accidents or incidents, or on similar cases that had been used in previous developmental tests or studies. Specific scenarios used in the experiment are summarized in the Appendix. The sequencing of scenarios within a crew’s set,

or across different crews, was randomized. However, the initial flights used for training and familiarization were the same for all (i.e. scenarios 2-1 and 2-3).



Figure 1. CMF/RFD Simulator (Note: For this study, the EFBs were re-positioned inboard and lower than shown; the right-side HUD is not shown.)



Figure 2. Head- and Eye-Tracking System Infrared Camera (inset) and Locations

III. Results

Over the course of the study, 10 two-pilot airline crews participated, with each crew spending two full days onsite and generally following the schedule given in Table 1. Each crew consisted of a Captain and First Officer from the same airline, with four different airlines represented. In total, 238 flights were completed, with flight times ranging between 8 to 15 minutes. A large body of data was recorded, and much of the analysis remains underway. In this paper, and in this section, selected preliminary findings and observations are presented related to the categories of complexity previously discussed in Section I. A more comprehensive report is planned for the future.

One important caveat to the results concerns the use of flight simulators by airline pilots. As this is the primary means by which they receive training (both initial and recurrent) from their airline employer, there are a set of biases they bring to the study. First, pilots are already familiar and proficient with most of the equipment and systems. This bias is leveraged to mitigate or lessen the learning curve effect on the results. Second, pilots expect certain types of simulator behaviors, scenarios, and limitations. Some of these are not “like the real aircraft” or “like the real operations.” This was mitigated by using check pilots who were also airline pilots, to help during the design and testing of each scenario. Then, during the preflight briefings, the experiment crews were informed of any relevant “sim-isms” found by the check pilots. Third, although many of the systems, equipment, and operations were the same as they use to when flying for their airline, many were very different and new. Fortunately, the crews adapted quickly and became somewhat proficient within only a few flights. Still, the findings and observations presented here are primarily drawn from the Day 2 flights so that learning curve effects for the new concepts should be less of a factor. In addition, this part of the study was not designed to evaluate pilot performance, rather to utilize experienced pilots to help discover complexity-related issues.

Table 2. Independent Variables for the Experiment

Independent Variable	Values
Weather conditions	(0) Cloud ceiling: 1000 ft, Visibility: 3 nmi, Dusk, Storm (None) (1) Cloud ceiling: 0500 ft, Visibility: 2400 ft, Night, Storm (Feb 25, 2011) (2) Cloud ceiling: 0200 ft, Visibility: 1800 ft, Night, Storm (Apr 27, 2011)
Airport configuration	(0) North flow (Runways 36L/C/R) (1) South flow (Runways 18L/C/R)
Traffic	(0) None (1) 3 nmi in-trail spacing, 1.5 nmi stagger, no departures (2) 5 nmi in-trail spacing, 2.5 nmi stagger, with departures on the outer runway
RNAV-based arrival and approach path	(FSD-18R, FSD-36L) Arriving from Northwest (EWR-18C, EWR-36C) Arriving from Northeast (MCO-18C, MCO-36C) Arriving from Southeast (IAH-18R, IAH-36L) Arriving from Southwest
Off-nominal condition	See Appendix
Crew	10 two-pilot crews

A. Complexity in Operations

Complexity in operations was represented by combining several aspects of operating concepts being promoted for the future with the current *modus operandi*. As described in Section I and the Appendix, fundamental elements of the new concepts were emulated and required changes to roles, responsibilities, level of automation support, and procedures. Our interest was not to study the effect of the added complexity of each of these operating concepts individually, but rather the combination of all (e.g. inter-dependencies and inter-operability).

Nine flights (scenarios 2-1(3) and 2-3(6)) were completed following initial training runs (i.e. familiarization flights). These did not involve complexity associated with the defined off-nominal conditions. The only differences were in-trail spacing (3 nmi vs. 5nmi), and airport configuration (North flow vs. South flow). These scenarios were briefed, and used, as the familiarization flights; however, performance was quite good for many of the nine flights that were completed afterwards (using the same scenarios). For example, for one of these flights, the pilot-flying began anticipating ASTAR-generated speeds by slowing early to smooth the turn onto final, but still achieving the spacing goal at the FAF. Also, some crews commented during these initial flights, as expected, that there was too much work involved with data link interaction on the EFB; however, in later flights the same crews felt data link interaction was not too demanding. During six of these nine flights, runway incursion warnings were triggered either due to the aircraft in front not clearing the runway in time (2 of 6), or due to the aircraft behind landing before ownship had exited (4 of 6), or both (2 of 6). For one of these flights, a traffic advisory was issued by the Traffic Alerting and Collision Avoidance System (TCAS) near the merge point, but resolved after the aircraft turned onto final. The crew noticed the advisory, but took no action.

All other flights (229) were based on manipulations of complexity using the independent variables given in Table 2, and prescribed by the scenario descriptions given in the Appendix. Based on all interviews, questionnaires, and data collected, the following list provides selected key observations from crews and the research team with respect to operational complexity. These observations are listed in no particular order.

1. It is critical to have well-defined procedures, in particular for FIM, DataComm, and use of EFBs. Further, designers must assure that procedures for one concept do not interfere or confound procedures for another.
2. With regard to FIM, well-defined rules, procedures, and training are essential regarding how fast to speed up or slow down in response to system-provided suggestions for changing speed. These rules, procedures, and training should include how or whether to use speed brakes and/or manage flap settings, at the very least. Procedures should never require flaps to move up after lowering, unless executing a missed approach.
3. There is some concern that FIM may lead to, or encourage, unstable approaches, particularly in lieu of all the other complexities involved with the approach and landing phases of flight.

4. There should be a speed limit defined at the FAF (e.g. ASTAR should never advise more than X knots); an alternative would be to require that the “achieve by” point is always at least Y nmi from the threshold. Appropriate values for X and Y would need to be determined, and based upon the category of aircraft involved and possibly other factors.
5. The runway incursion indications and alerts are very good, except when they are triggered by the aircraft landing behind; notification in these cases should be much less salient to avoid the risk of encouraging “hurrying” off the runway (e.g. no red, no aural). As implemented here, there were cases where crews felt added pressure to get off the runway faster than they normally would. Some crews did not want to know about the aircraft behind; some did want to know.
6. Procedures should allow for pilots to continue an approach to landing even if another aircraft is on the runway at the same time (this happens frequently in real operations), but only if cleared by ATC, and if there is visual indication (outside or on a display) that loss of separation on the runway will not occur (e.g. the lead aircraft is clearing the runway).
7. Multiple crews recommended putting EFB(s) in Flight Operations areas so that crews can use them to preview and plan their flights; then, when in the cockpit, the interface will be familiar and they can review the same information, or check for updates.
8. The communications environment should be all voice, all data link, or voice-by-exception; and this may vary based on phase of flight, or by operating area. Regardless, training and procedures are critical to any change. Voice-by-exception was recommended by all crews; noting that mixing modes, as was done in this simulation, can be confusing, particularly if the voice and data messages disagree. If a mixed-mode environment is used, crews recommended removing FIM-related details from the voice instructions; also suggesting that some clearances be voice only (e.g. “cleared to land”). The system should not send long text messages (ATC or other); perhaps limiting length in a manner similar to cell phone-based text messaging.
9. EFVS was viewed as good support for EVO; and all crews believed it could be effective to a 50 ft decision height, or less. However, some of the crews had very few low visibility approach/landing experiences, other than in flight simulators.
10. Several crews recommended voice-only communications below a defined altitude (e.g. 3000 ft or 5000 ft); but allowing DataComm on the surface after landing, or before takeoff.
11. Any weather-related information updates (e.g. wind, runway conditions) should be provided via voice when aircraft are below 3000 ft on approach.
12. There is risk of loss of critical elements of situation awareness in a data link-only environment; crews benefit from hearing what others are cleared to do.

B. Complexity in Systems and Procedures

Based on all interviews, questionnaires, and data collected, following are selected key observations from crews and the research team with respect to systems and procedural complexity. These are listed in no particular order.

1. Closely-spaced parallel approach operations when combined with FIM and EVO will likely require roll-out guidance to an appropriate exit; without this capability, crews will feel pressure to exit the runway prior to the trailing aircraft landing. This combination of concepts also requires an additional system to detect and warn of intruders (beyond current TCAS capability); the observed occurrences of apparent Inattention Blindness (IB)¹⁹ implies that pilot monitoring of a traffic display is insufficient for reliable detection. See Section III.E for more information on IB.
2. The DataComm message scratchpad is recommended, as messages are always in view, and there is a reduced likelihood of forgetting or making a mistake regarding ATC instructions. However, crews tended to believe these messages were always correct, even when the voice or graphically depicted instructions differed. The procedure for resolving such differences must be well defined.
3. System designers must assure that critical information is not occluded. This occlusion issue was most noticed here when the weather radar was used in conjunction with the traffic display and FIM status on the ND. There were also circumstances that caused information occlusion on the EFB.

4. Some heuristics should be established for the magnitude of incremental changes to ASTAR-provided speed changes. For example, no greater than +/- 10 knots change when nearing the FAF.
5. The surface moving map should be available to both crewmembers; preferably upfront and on an automatic switch (e.g. weight-on-wheels) that can be over-ridden. All crews felt the surface map function was excellent and much needed. One crew suggested de-emphasizing surface traffic when in flight (e.g. use dot symbols).
6. Although the EFB was judged easy to learn and use, a few shortcomings were identified. (a) An efficient method is needed to interact with the data link messages, while also allowing a quick return to the relevant chart. Crews felt the relevant chart should be displayed by default at all times, with the pilot able to override to review/respond to a data link message; or view other information, but able to quickly return to the chart. One idea was to have EFB automatically return to previous page following "wilco" of a data link message. (b) Charts should sequence automatically (Arrival-Approach-Airport), or via a single button press. (c) Crews would like to be able to build/review the flight plan in the EFB. Further, the EFB should be connected to the flight management system to obviate the complexity and difficulty of building flight plans in traditional control display units.
7. The FIM-related PFD symbology was commended, encouraging attentiveness and engagement by being spatially separated and combined with an audible tone announcing a change.
8. All crews viewed the display of NOTAMs, both for pre-flight and as updated in-flight, as a significant improvement over current state-of-the-practice. Several crews recommended sending relevant AIS/MET data link messages directly to the NOTAM pages in the EFB, rather than the intermediate communications pages. Also, when there are multiple closures, these should be provided as one comprehensive NOTAM (e.g. "Airport closures, see chart").
9. For the FIM-related displays, it would help to see information about the aircraft behind. This could allow the crew to consider taking an early exit prior to landing.
10. If replacing voice communications with DataComm, it was suggested to show pilots what other aircraft are cleared to do.
11. There is a need for an improved way to manage range setting on the ND when switching from the in-flight mode to the surface map mode. When approaching a runway, or during a go-around, the range setting on the ND will be "zoomed in"; but after landing, and switching to surface map mode, the range setting needs to be "zoomed out". This may require separate controls for setting range.
12. All crews commented favorably on the utility of the runway inset display. This display allowed them to preview the runway environment prior to landing, check for traffic on the runway, and preview the planned exit and taxi path. This observation reinforces previous findings.

In terms of procedural complexity, the following behaviors were observed. Three crews decided to go slower than ASTAR-provided guidance near the merge point (200 knots instead of 220/230 knots). Multiple crews anticipated what the ASTAR algorithm would suggest for the speed change. Their "guess" was not always correct. One crew executed a small 'S' turn to slow down and increase spacing when ahead of the FIM spacing goal. One crew seemed to think the ASTAR-provided speeds were coming from ATC.

C. Complexity in the Flight Environment

With respect to weather-induced complexity, seven flights (scenarios 2-6(4), and 2-8(3)) were completed wherein runway conditions were degraded and reported via either ATIS or the AIS/MET data link service. For these flights, the two storm events were represented, along with the commensurate weather radar and EFB weather displays. Turbulence was light to moderate. For the four flights involving icy runway conditions, three landings were completed without exiting the runway (stopping on the runway). The other ended in a go-around due to no lights visible at the decision altitude. For all such flights, the crews stated that in real operations they would expect a pilot report from leading aircraft via voice radio. For the three flights involving wet runway conditions (reported as "poor"), the crews consulted the aircraft operating manual to determine runway landing length required, then continued the landing and exit without incident.

Off-nominal aircraft conditions included: (1) a failure of the antiskid system, (2) a loss of hydraulic pressure, and (3) a loss of the ILS signals. Twenty-one flights involving onboard system failures were completed. The failure of the antiskid system was noticed in all 11 flights that it occurred, with a longest time-to-notice of about 3 minutes.

For 10 of these 11 flights, the crew worked through the appropriate procedure/checklist in addition to the other procedures and tasks required, and continued the flight. For one case, the crew did not perform the checklist, merely commenting on the antiskid indication. Four of the 11 flights ended in a go-around; the other seven were completed after landing and runway exit. Two of the go-arounds were declared after the pilots discussed the failure and the runway conditions/length, then decided there was insufficient runway length to continue. The other two go-arounds were due to more insidious conditions: (1) the crew forgot to set the ILS frequency, and (2) the crew forgot to engage the approach mode in the autopilot. The latter case resulted in the aircraft staying high (4000 ft MSL) to the FAF, where the crew first detected this inappropriate altitude – an unstable approach event. However, a go-around was declared and flown without incident from this point. For one of the successful landings, the crew decided to engage autoland. For all others, the autopilot was disengaged at points after the FAF. For each instance of the antiskid system failure, there was good discussion by the crew of whether to proceed. In one case, the crew felt that in normal operations they would not continue the flight given the ATIS-reported conditions (as provided by the EFB). Runway incursion alerts were triggered for two of the seven completed landings, both caused by the aircraft landing behind before the subject crew had cleared the runway.

The second emulated system failure was a loss of hydraulic pressure occurring early in the flight. This condition was noticed by the crew in all 10 flights that it occurred. However, due to the nature of this failure and the added and various complexities (e.g. FIM and weather), crew reactions were more diverse than for the antiskid system failure. All crews worked through the emergency checklist, but several were interrupted by other tasks (e.g. achieving ASTAR-suggested speeds). In addition, some crews contacted ATC and declared an emergency, then broke out of the traffic pattern to work the issue. In addition, other conditions occurred for some of the flights coincident with the system failure. For example: (1) three of 10 flights received a TCAS traffic advisory near the merge point that resolved after the aircraft turned onto final, (2) four of 10 flights included runway incursion detections of other aircraft on the runway either before or after landing, and (3) one of the flights was broken off after receiving a windshear warning via the AIS/MET data link service. Only five of 10 flights were continued through landing and runway exit. For one of these cases, the crew made a wrong turn during taxi.

The ILS failure condition only occurred on three flights. For each of these cases, the crew detected the problem quickly and initiated an immediate go-around. The added complexities appeared to have no effect on decision making with respect to this condition, largely due to the salience of the failure indication (on the PFD) and the airline-trained procedure for taking action.

D. Complexity Due to Uncertainty

Sixteen flights of (scenarios 2-5(11) and 2-12(5)) were completed wherein a leading aircraft along the approach path slowed down while approaching the runway creating a logjam of traffic, all slowing down to maintain the desired interval. For these flights, the test crew was asked by ASTAR to slow to an airspeed as low as 130 knots just after the merge point in order to achieve the spacing goal from the slower aircraft in front. Weather was not an added complexity or variable in these flights; however, for eight of the flights, AIS/MET data link messages were sent just after the start of the flight regarding a closure of the last 500 ft of the landing runway, and a change of the decision altitude for the approach procedure. Neither was operationally significant, but needed to be considered. These messages were not sent for the remaining eight flights. For all flights, the crew was able to reduce speed as suggested by ASTAR. Maintaining spacing was more challenging given low energy and an early landing configuration in terms of flaps and gear (down). For three flights, the in-trail spacing got lower than the required 2.0 nmi (as low as 1.5 nmi), and went unnoticed by the crews. For two flights, the aircraft arrived at the FAF early (10-20 sec). For one flight, the aircraft arrived excessively early to the merge point resulting in an abrupt ASTAR-based speed change recommendation, from 200 to 130 knots. Pilots commented that an instantaneous change this large should not be suggested by the system this early in the approach. Pilots also commented that the indication of status with respect to achieving the FIM goal at the FAF should be represented in three forms: time (sec), speed (knots), and distance (nmi). Six of the 16 flights ended in a go-around, declared very late in the flight (as low as 150 ft). These were all due to the aircraft in front not clearing the runway prior to landing. In each of these cases, runway incursion alerts were provided, but the lead aircraft was very close to exiting. Pilots commented that such situations occur in real operations and often ATC will anticipate the exit and allow the aircraft to land.

Twenty-nine flights (scenarios 2-16(11), 2-17(11), and 2-22(7)) were completed involving an aircraft intruding into or across the flight path of ownship. All intrusions occurred after ownship had passed the merge point and was on runway heading. For 22 of the flights, the aircraft just ahead and going to the parallel runway began a slowly developing turn onto ownship's track as if the intruder were following the ILS to the wrong (ownship's) runway, or had selected the wrong runway in the FMS. For 12 of these 22 flights, the crews noticed the intruder and took action; either by working with ATC to resolve the problem, by breaking out of the traffic pattern, or by monitoring

and maintaining spacing with the intruder. However, for 10 of these 22 flights, the crews never noticed, or commented on, the intruder aircraft. Several different behaviors resulted. In nine cases, crews continued to maintain spacing with the FIM-paired aircraft (in front of the intruder) and followed ASTAR suggestions for speed adjustments. In one case, the crew confused the intruder aircraft with the FIM-paired aircraft (ahead of the intruder). For all 10 of the “never noticed” cases, it was only after the intruder had landed and the crews realized the runway would not be clear and therefore, a go-around was required. In one case, the crew commented that they felt the aircraft on the runway was a departing aircraft that had been mixed in by ATC. All of these “never noticed” cases are believed to be examples of Inattentional Blindness (IB)¹⁹ – to be discussed further in Section III.E.

For seven of the 29 flights, the aircraft going to the parallel runway, and just ahead, began a slowly developing turn across ownship’s track, executing a 30-degree bank angle turn and climbout, as if the intruder aircraft were initiating a go-around in the wrong direction - toward ownship’s runway. Only one of the seven occurrences was noticed by the crew prior to crossing the flight path. A second was noticed after the intruder had completely crossed the approach path. The remaining five flights were completed without incident; however, as with above, it is believed that these may be cases of IB.

Sixteen flights (scenarios 2-19(13) and 2-21(3)) involved a runway incursion wherein an aircraft on the surface taxied into position-and-hold (as if preparing to depart) on the runway to which ownship was approaching. The aircraft began moving into this position just after the lead (FIM-paired) aircraft had landed. In all cases, the crew executed a go-around after receiving the runway incursion warnings regarding the runway traffic. One crew commented that under DataComm operations (such as employed in this study), pilots on approach would not hear ATC clear the incurring aircraft into position-and-hold, making the runway incursion detection and alerting functions even more critical.

E. The Manifestations of Complexity

In Section I.E a set of known manifestations of complexity were given. Although rare in airline operations, it was expected that they would occur during this study given the circumstances encountered through the design of the test scenarios. A summary of key observations regarding each type is given here.

Because only approaches were conducted, there were no cases of runway incursion by the test crews during taxi; however, there were several cases where crews were slow to exit prior to the aircraft behind landing. Given the tight in-trail spacing, this was an expected result, and may be expected in future operations resulting in the potential for increased go-arounds. One mitigation strategy would be to provide roll-out guidance to high-speed exits. There were a small number of runway excursions due to runway conditions (wet or icy). Each of these was most likely due to a limitation of the simulator wherein the runway surface appeared visually (out the window) to be “dry,” and counter to the reported runway conditions per ATIS and/or AIS/MET data link provision on the EFB. For most flights involving wet/icy runways, the crews managed landing roll-out well and likely better than they would have if runway conditions were not reported (via the EFB). As expected, the tailwind scenarios did not induce runway excursions, primarily due to more-than-adequate runway length and use of autobrakes.

There were no observed instances during flight of spatial disorientation (i.e. loss of attitude awareness), wake encounters, or inadvertent path deviations. It is possible a wake encounter may have occurred for the cases of loss of in-trail spacing (as low as 1.2 nmi); however, wake turbulence and its effects were not modeled. Multiple crews did comment that FIM may increase risk of wake encounters if interval goals are set too small. Path deviations did occur on the airport surface in a few instances, including (a) exiting the runway onto a closed taxiway, and (b) making a wrong turn after exiting. The first deviation was due to the crew’s desire to exit early so that the trailing aircraft would not have to perform a go-around; the second deviation was due to distraction while the crew discussed the landing. All crews felt that the moving map combined with the EFB-based mechanism for providing NOTAMs would significantly reduce the likelihood of unsafe path deviations on the airport surface. There was one case in flight where the Captain and First Officer were referring to two different approach charts (Runway 36L vs. Runway 36C) on their EFBs resulting in confusion and some debate as to the correct procedure and frequencies to use. After discussion, the discrepancy was resolved; however, this issue should be addressed when designing procedures for EFB use in the future (e.g. add to the approach checklist to cross-check the selected landing runway in the EFB).

Unstable approaches (e.g. loss of energy state awareness) and automation mode confusion situations occurred for a small number of flights. This is partially due to the learning curve associated with the new concepts and the facility. However, some cases require additional investigation as to cause. For example, two crews were unaware that they had failed to engage the approach mode in the autopilot, keeping them above the glideslope (in one case at 4000 ft at the FAF). There were other cases where VNAV was engaged/disengaged by the crew with some frustration, and lack of knowledge, as to why there was difficulty for the system to capture the vertical path. There

was also some occasional confusion resulting from interacting with the FIM-related automation mode (called Paired Dependent Speed (PDS)). Analysis of results regarding these issues continues, and will be reported in a future paper.

Loss of separation did occur, but most were during scenarios designed for this purpose. The goal here was to observe whether the crews noticed these situations, and if so, what decisions resulted. As described previously, several cases of apparent IB occurred. Of the 29 flights involving an intruder aircraft, 16 of the intrusions were not noticed by the crews. This despite both crewmembers having access to a traffic display, being aware of ATC-provided no-closer-than instructions, and being paired with a lead aircraft per the FIM operation. IB is a failure to notice an unexpected cue in the field of vision when other attention-demanding tasks are being performed¹⁹. This has been shown to occur when humans are overloaded with visual stimuli (i.e. it becomes impossible to pay attention to all cues), and they are unaware of the unattended cue. There is a significant amount of further data analysis required to confirm this condition, but the recorded head and eye movements should enable a conclusive finding.

F. Summary Observations Regarding Decision Making

One item on the post-flight questionnaire asked pilots to rank the factors that most influenced their decisions with respect to maintaining safety during the flight. Each pilot was asked to list at least a Top-3 after each flight. The aggregate result in terms of a Top-10 was: (1) traffic, (2) navigation, (3) communication, (4) energy management, (5) autopilot mode, (6) aircraft systems state, (7) weather, (8) aircraft attitude, (9) NOTAMs, and (10) flight planning. The two most explicit examples of overall complexity affecting crew decision making were: (a) when/how to use automation, and (b) when/if to declare a go-around, or break out of the traffic pattern.

Regarding the use of automation, a distinction must be drawn between control automation and information automation (i.e. decision support). Decision making with respect to the use of control automation (e.g. enabling modes such as LNAV and VNAV) generally followed airline procedures as supplemented by the study-specific training. Crews were given the option to disengage the autopilot at the FAF and hand-fly the landing, and this was done for most flights when weather conditions allowed. Four flights involved errors wherein the crew either failed to engage the approach mode, or failed to tune to the correct ILS frequency. In terms of information automation, decision support was judged excellent; however, there was an implied and significant reliance on all such support to be available and trustworthy (i.e. correct). For example: there was a reliance on TCAS to protect against loss of separation; there was reliance on ATC and/or airline dispatchers to provide support, or intervene, as required; and there was a reliance on the information provided by the displays to be complete, accurate, up-to-date, and of high integrity. In some cases when these properties did not exist, poor decisions resulted.

Decisions to go-around or break out of the traffic pattern occurred relatively frequently (58 flights, or ~25%). Nineteen of these decisions were expected per the scenario design (e.g. runway incursions and ILS failures) and were made by all crews consistently. There were several reasons for the remaining 39 flights, including: the lead aircraft not quite clearing the runway (16), working to resolve a system failure (7), loss of in-trail separation (7), unable to see the runway environment at the decision altitude (5), unstable conditions (3), and receiving a TCAS traffic advisory (1). Each was due in large part to the complexity encountered and may not have been necessary if some complexities had been removed.

In summary, although the many disparate concepts and technologies employed in this study are at various stages of technological or implementation readiness levels, all of the crews coalesced on a largely consistent work environment and set of procedures. Despite limited training time, none commented on being overwhelmed, or task saturated in general, with only a few exceptions for short periods of time. Further, when asked during the summary debrief whether they felt all of this was “doable”, 11 of 12 crews said yes, and that by and large, the decision support was far better than they have today. However, an over-arching concern was expressed regarding the criticality of defining effective procedures and training protocols. The more the deviation from current practices, the more critical this factor, and the larger the risk of new forms of decision-making errors and unsafe situations.

IV. Conclusions and Next Steps

This paper describes a flight simulation study designed to collect data to help in understanding how complexity can affect crew decision making. Flight simulation is used extensively by airline pilots as a tool to learn. Here experienced crews were asked to learn, and then to help the research team discover issues related to decision making and decision support that may require further research and/or a re-definition of designs or operating concepts. The experiment was conducted to help answer a very broad and difficult research question. Given the totality of the complexities that flight crews are expected to deal with both now and in the future, and the requisite training and transition processes, will the risk of unsafe situations reduce, increase, or stay the same? While this single

experiment cannot conclusively answer such a broad question, it can and does provide evidence of movement toward a level of complexity that can induce new decision-making behaviors (e.g. more go-arounds) and risks (e.g. unrecognized loss of separation).

As mentioned previously, one property of complexity is “difficult to understand.” As such, learning and experiencing an operational environment can be, in its own right, a means to remove complexity. However, a second property “made up of many inter-related parts” is not affected by learning; in fact, the number of inter-related parts that crews will have to be aware of will continue to grow according to plans for the future air transportation system. Interestingly in this study, although all crews acclimated quickly to the simulated environment and felt the transformations they experienced could improve efficiency and safety, there were several behaviors and decisions observed during these flights that should be considered moving forward. The most important of which may be the effect that the inter-relationships and inter-operability of the many operational improvements being planned and developed can have. Advancing each improvement along an independent thread may lead to unexpected and added complexities that cannot be removed (e.g. by training) when later combining with others. A second, and perhaps equally important, observation involves the reliance on control and information automation. Such a reliance can only improve decision making if (1) the information provided is complete, accurate, up-to-date, and of high integrity, and (2) the crew understands the limits of what the automation can and cannot do (i.e. its intended function).

The next steps for this research are in the near-term to: (1) complete the analysis of the head and eye tracking data (e.g. to determine how much attention is being taken by EFBs, and to confirm the IB hypothesis), (2) report on the usability assessment of the EFB for the NOTAM and data link applications, and (3) complete the analysis of data regarding the use and awareness of automation states. In the longer term, new experiments are planned to find ways to remove system complexity and manage information such that improved decisions result, particularly in situations where attitude, energy, or system state awareness is critical. Loss of awareness of these three states has been a significant contributing or causal factor in multiple recent accidents, and as a result, research to develop interventions strategies has been recommended.

Appendix

Figures A-1 to A-10 illustrate the flight deck displays used for this study. Table A-1 summarizes the scenarios used during Part 2 of the experiment.



Figure A-1. CMF/RFD Display Configuration.



Left EFB



Right EFB

Pilots access sub-menus by pressing touchscreen label or the associated bezel button



- COMM – To review and acknowledge datalink messages from ATC and the AIS/MET service
- WEATHER – To review MET reports (METARs), Terminal Area Forecasts (TAFs), Airmen’s MET reports (AIRMETs), Pilot reports (PIREPs), Winds aloft, and composite ground-based radar images
- NOTAMS – To review NOTAMS in text and graphic forms (during pre-flight or when received in-flight via the AIS/MET service)
- AIRPORT MOVING MAP – To view airport features with position overlay, traffic, and taxi route
- CHARTS (Arrival, Approach, Airport)
- CHECKLISTS (Descent, Landing, Taxi)
- ATIS (Automatic Terminal Info Service)

Figure A-2a. EFB Main Menu.

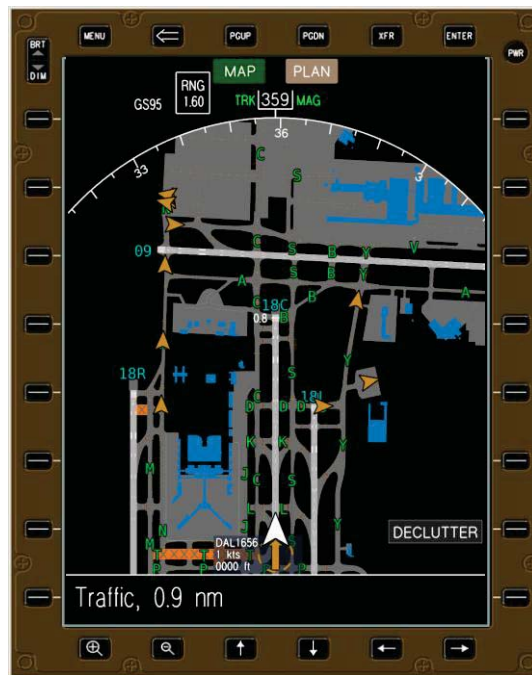
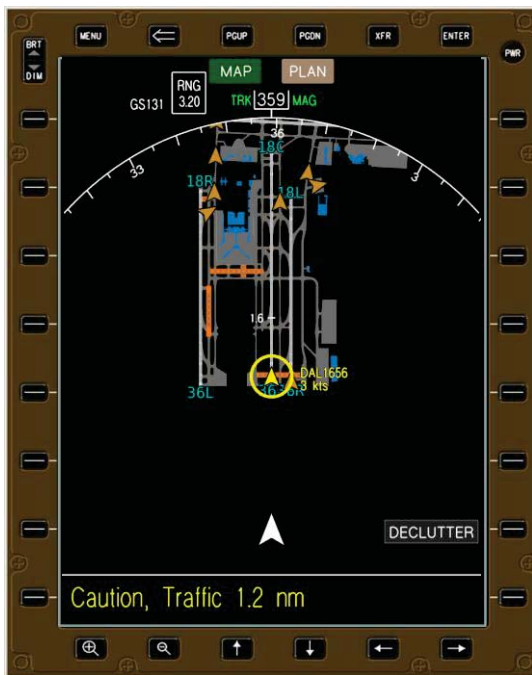


Figure A-2b. EFB Airport Surface Moving Map Page (left – on approach; right – on the surface)

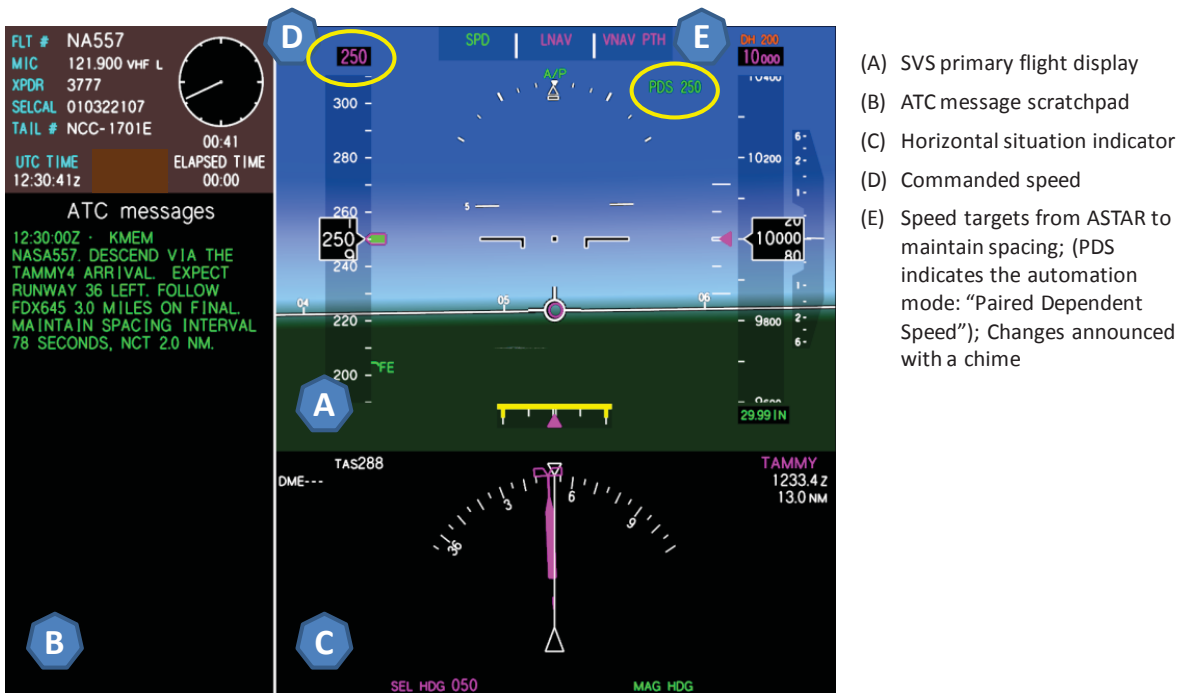


Figure A-3. Left-Side Outboard Display.

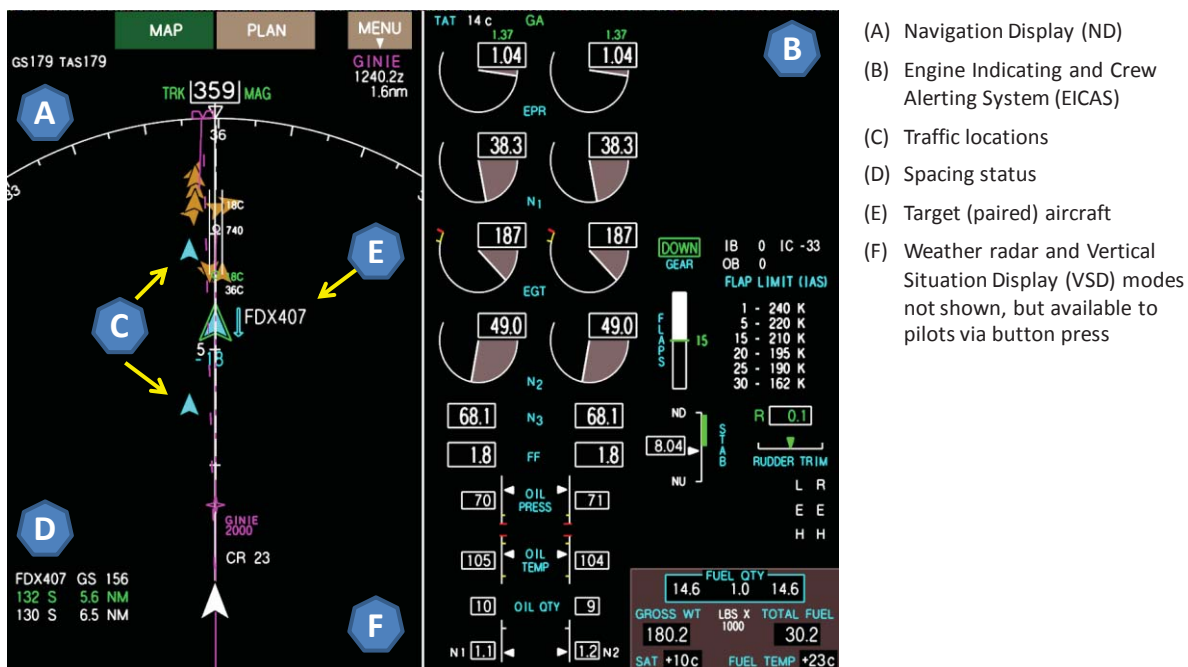
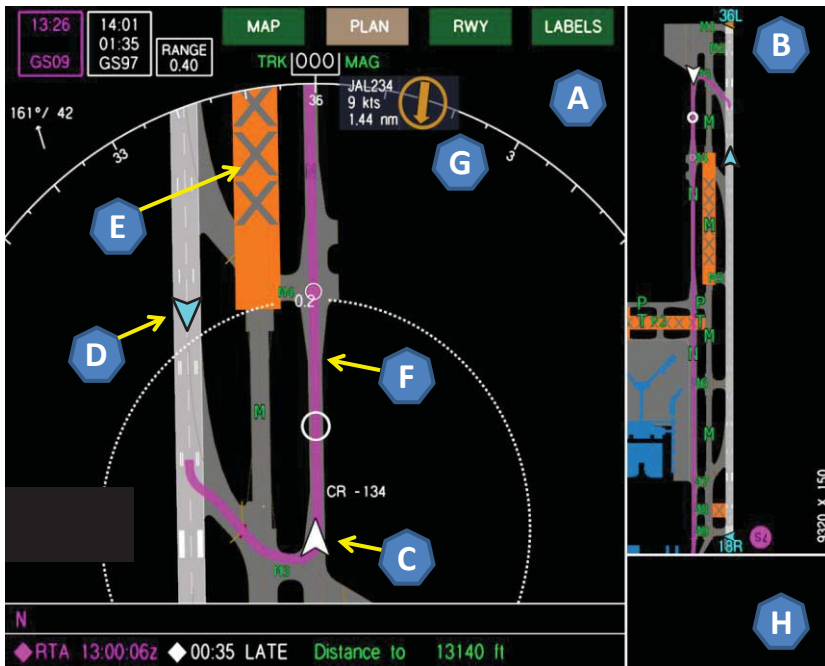


Figure A-4. Left-Side Inboard Display.



- (A) Airport surface moving map
- (B) Runway inset
- (C) Current aircraft position
- (D) Traffic locations
- (E) Taxiway closure (NOTAM)
- (F) Taxi route from ATC
- (G) Surface traffic off-scale
- (H) Pilots can switch to ND/EICAS as shown in Fig. A-4 via button press

Not shown: Runway incursion indications (runway outline turns blue, yellow, or red)

Figure A-5. Right-Side Inboard Display.

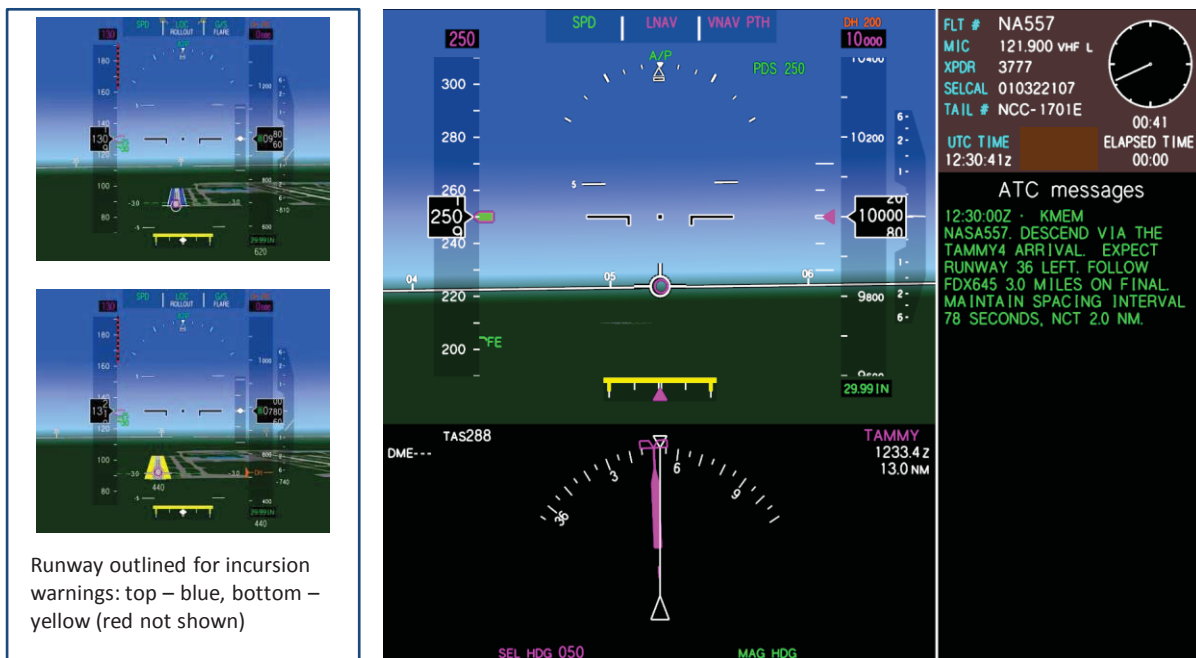


Figure A-6. Right-Side Outboard Display (right); Examples of Incursion Warnings on PFD (left).

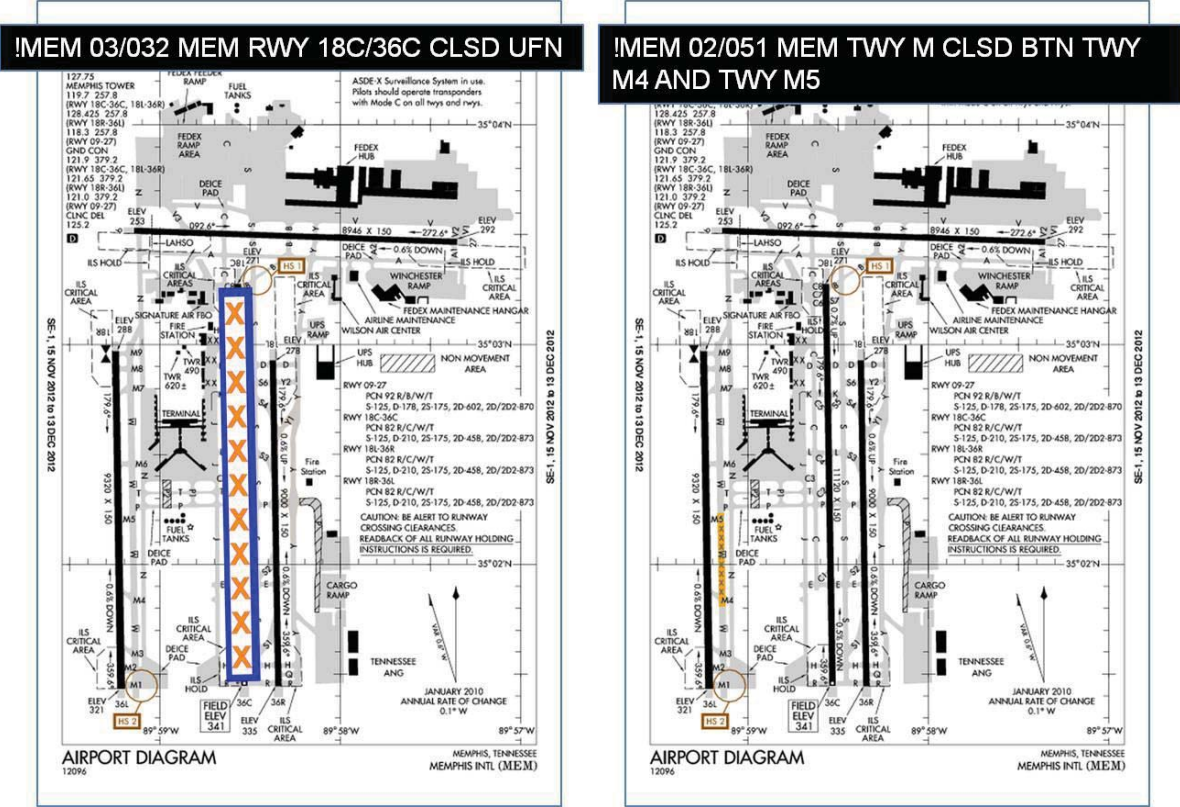


Figure A-7. Example Graphical NOTAMs as displayed on EFB.



Figure A-8. Synthetic Vision PFD at Multiple Altitudes During an Approach.

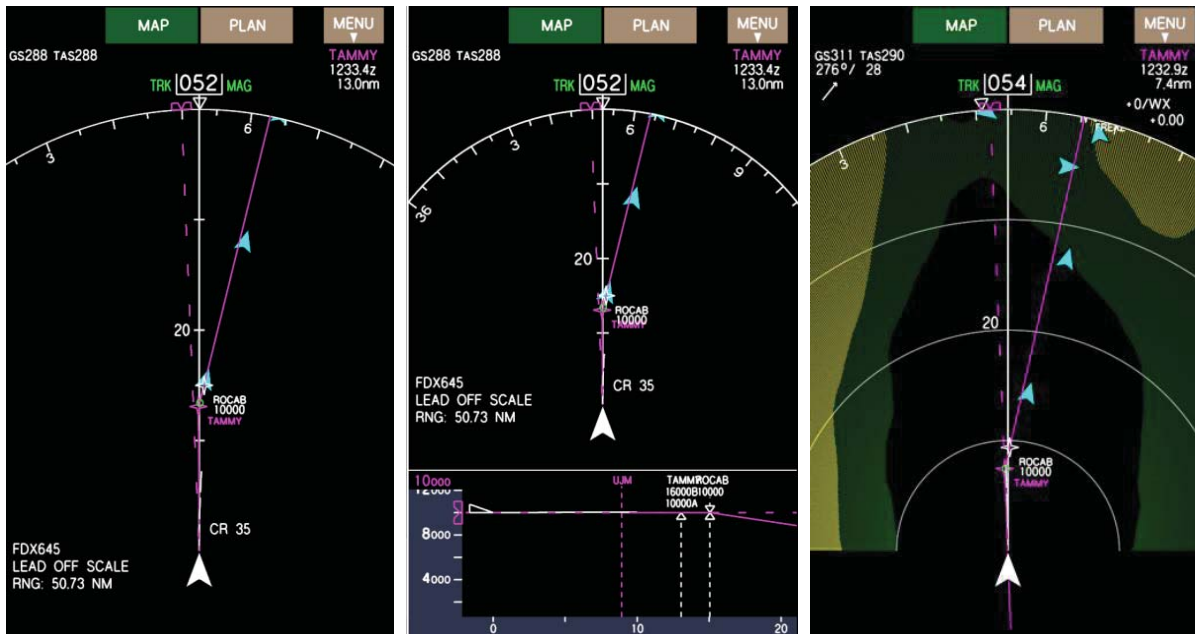


Figure A-9. Navigation Display with Examples of VSD ‘on’ (center) and Weather Radar ‘on’ (right).

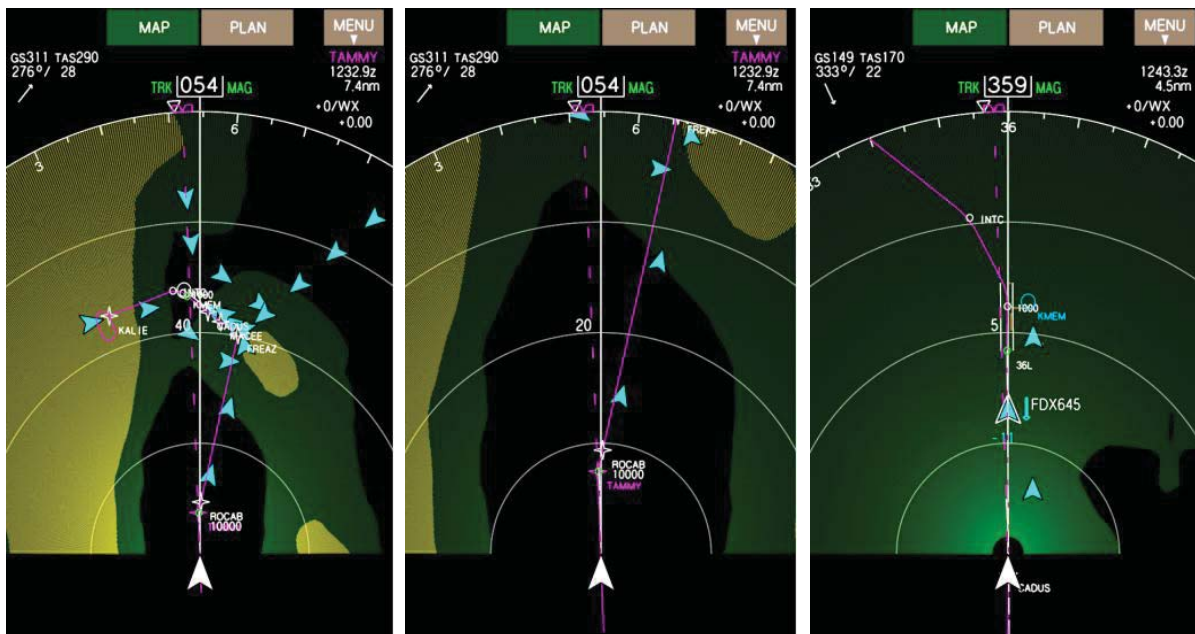


Figure A-10. Navigation Display with Examples of Weather Radar at Different Range Settings.

Table A-1. Summary of Part 2 Test Scenarios.

No.	Wx Case	Airport Config.	Traffic Case	Path	Off-nominal
2-1	0	1	1	IAH-18R	None.
2-3	0	0	2	MCO-36C	None.
2-4	2	0	1	IAH-36L	AIS/MET DL message sent at 7500 ft: KMEM temporary obstruction (crane). ANTISKID warning on EICAS 30 sec after start.
2-5	0	0	1	EWR-36C	AIS/MET DL messages sent 1 min after start: KMEM Rwy 36L north 500 ft closed, and change to IAP for ILS or LOC Rwy 36C (DA 551 ft, HAT 230 ft, RVR 2400 ft). Traffic in front slowing.
2-6	2	0	2	FSD-36C	AIS/MET DL message sent near FAF: Current threshold winds from 180 deg at 10 knots. Freezing rain/drizzle reported in ATIS. Braking performance on rollout as if freezing rain.
2-7	0	1	1	MCO-18C	AIS/MET DL message sent at 3000 ft: Low level wind shear, +/- 10 knots, 1000 ft on approach Rwy 18C. EICAS message – “HYD SYS PRESS (CENTER ONLY)” at 30 sec after start.
2-8	1	1	1	EWR-18C	AIS/MET DL message sent at 5000 ft: KMEM Rwy 18C braking conditions poor.
2-9	1	1	1	IAH-18R	The aircraft sequenced to land in front has ADS-B performance is reduced such that incursion detection algorithms will not function. Aircraft will still be displayed, but no alerts are provided regarding runway occupancy.
2-10	0	1	1	FSD-18R	Loss of ILS signal at 1200 ft.
2-11	2	0	1	FSD-36L	Reduced braking on landing and rollout as if heavy rain.
2-12	0	0	2	MCO-36C	Traffic in front slowing condition.
2-16	0	1	1	EWR-18C	After passing the merge point and on runway heading, the aircraft going to the parallel runway, and just ahead begins a turn toward own-ship’s track - a slowly developing turn as if the intruder begins to follow the ILS to the wrong runway.
2-17	1	1	1	IAH-18R	When approaching the merge point, the aircraft to be sequenced in front on the parallel runway approaches the merge point as if heading to the wrong runway - emulating the intruder’s crew selecting the wrong runway in the FMS.
2-19	2	0	2	EWR-36C	Runway incursion event – aircraft moves onto runway as if to depart, just after IM-paired aircraft lands.
2-21	2	0	2	EWR-36C	Runway incursion event – aircraft moves onto runway as if to depart, just after IM-paired aircraft lands. ADS-B performance is reduced.
2-22	1	1	1	FSD-18R	After passing the merge point and on runway heading, the aircraft going to the parallel runway, and just ahead begins a turn toward own-ship’s track – a 30 deg bank angle turn with a climb initiated, as if the intruder aircraft is initiating a go-around in the wrong direction.

Acknowledgments

The authors would like to thank the following individuals who were key contributors to the design and execution of this study: Capt. Daniel Kiggins (American Airlines), First Officer Marc Henegar (Alaska Airlines and ALPA), Tim Etherington (Rockwell-Collins, Inc.), Mike Norman (Boeing), and Kathryn Hoekstra (FAA). In addition, collaborators from NOAA provided unique weather models and expertise that were essential to project success. Thanks to Bryan Barmore and Terry Abbott for their help with FIM and ASTAR. Last, thanks to NASA LaRC's simulation development team, led by Victoria Chung and James Barnes, for their creativity and dedication. Turns out the study of increasing complexity requires a commensurate amount of increasing complexity be folded into a flight simulation environment; a challenge in its own right, to which this team took on, and overcame, beyond expectations.

References

- ¹Belcastro, C.M., and Foster, J.V., Aircraft Loss-of-Control Accident Analysis, Proceedings of the AIAA Guidance, Navigation, and Control Conference, Toronto, Canada, August 2-5, 2010.
- ²NextGen Implementation Plan, U.S. Department of Transportation, FAA, www.faa.gov/nextgen, June 2013.
- ³Prinzel, L. J., et. al., Flight Deck-Based Delegated Separation: Evaluation of an On-board Interval Management System with Synthetic and Enhanced Vision Technology, Proceedings of the 30th AIAA/IEEE Digital Avionics Systems Conference, Seattle, WA, October 2011.
- ⁴Safety, Performance, and Interoperability Requirements Document for Airborne Spacing – Flight Deck Interval Management, RTCA Document DO-328, RTCA, June 22, 2011.
- ⁵Barmore, B., et. al., A Concept for Airborne Precision Spacing for Dependent Parallel Approaches, NASA TM- 2012-217346, NASA, March 2012.
- ⁶Barmore, B.E., Baxley, B.T., and Murdoch, J., Research Of Airborne Precision Spacing to Improve Airport Arrival Operations, Air Traffic Technology International Magazine, pp 70-73, January 2011.
- ⁷Advanced Surface Movement Guidance and Control Systems Manual, ICAO Document 9830/AN52, ICAO, 2004.
- ⁸Young, S.D., and Jones, D.R., Runway Incursion Prevention Using an Advanced Surface Movement Guidance And Control System (A-SMGCS),” Proceedings of the 19th AIAA/IEEE Digital Avionics Systems Conference, Philadelphia, PA, October 7-13, 2000.
- ⁹Young, S.D., and Jones, D.R., Runway Incursion Prevention: A Technology Solution, Proceedings of the Joint Meeting of the Flight Safety Foundation's 54th Annual International Air Safety Seminar, the International Federation of Airworthiness' 31st International Conference, and the International Air Transport Association, Athens, Greece, November 5-8, 2001.
- ¹⁰Safety, Performance and Interoperability Requirements Document for Enhanced Traffic Situational Awareness on the Airport Surface with Indications and Alerts (SURF IA), RTCA Document DO-323, RTCA, December 2010.
- ¹¹www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/atc_comms_services/datacomm/general
- ¹²Minimum Aviation System Performance Standards (MASPS) for Enhanced Vision Systems, Synthetic Vision Systems, Combined Vision Systems and Enhanced Flight Vision Systems, RTCA Document DO-315B, RTCA, June 2011.
- ¹³Pilots Bill of Rights, Amendment to Title 49 of the U.S. Code of Federal Regulations, U.S. Congress, January 3, 2012.
- ¹⁴Concept of Use for Aeronautical Information Services (AIS) and Meteorological (MET) Data Link Services, RTCA Document DO-340, RTCA, September 2012.
- ¹⁵Abbott, T. S., “A Revised Trajectory Algorithm to Support En Route and Terminal Area Self-Spacing Concepts,” NASA/CR-2010-216204, NASA Langley, Hampton, VA, Feb. 2010.
- ¹⁶Murphy, K., The Paperless Cockpit, New York Times, p. B-6, July 5, 2011.
- ¹⁷Duan, P., et. al., A Simulation Environment For Evaluation of Integrated Alerting and Notification (IAN) Concepts, Proceedings of the 31st AIAA/IEEE Digital Avionics Systems Conference, Williamsburg, VA, Oct 14-18, 2012.
- ¹⁸Daniels, T. D., et. al., Flying A Motion-Based Piloted Flight Simulator Through Numerical Weather Models, Proceedings of the 31st AIAA/IEEE Digital Avionics Systems Conference, Williamsburg, VA, Oct 14-18, 2012.
- ¹⁹Mack, A., and Rock, I., Inattention Blindness, MIT Press, Cambridge, MA, 1998.