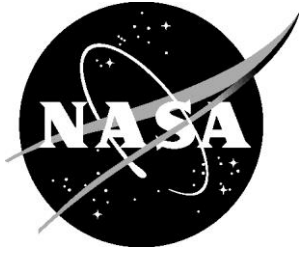# Avionics Configuration Assessment for Flightdeck Interval Management: A Comparison of Avionics and Notification Methods

Kara A. Latorella
Langley Research Center, Hampton, Virginia

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

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

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

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

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

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

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

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

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

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199
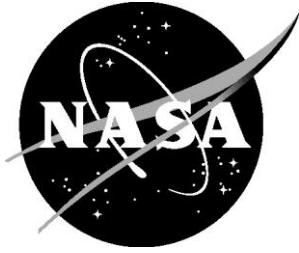
Avionics Configuration Assessment for Flightdeck Interval Management: A Comparison of Avionics and Notification Methods

Kara A. Latorella
Langley Research Center, Hampton, Virginia

August 2015

Acknowledgments

Table of Contents

List of Figures

List of Tables

Abbreviations

AAA	Notification Method with aural for new speeds, speed deviations, reminders
AC	(FAA) Advisory Circular
ADS-B	Automatic Dependent Surveillance-Broadcast
AGD	ADS-B Guidance Display
AGL	Altitude above Ground Level
AIAA	American Institute of Aeronautics and Astronautics
ANG-C1	Organizational code for FAA's Human Factors and Engineering Division
ANOVA	Analysis of Variance
AOI	Areas of Interest defined for oculometer analysis
ASTAR	Airborne Spacing for Terminal Area Routes
ATD-1	Air traffic management Technology Demonstration-1
ATIS	Automated Terminal Information Service
CDTI	Cockpit Display of Traffic Information
CONOPS	Concept of Operations
CR	Contractor Report
CRM	Cockpit Resource Management
dBa	Decibels – frequency-weighted for human aural perception
DASC	Digital Avionics Systems Conference
DFW	Dallas Fort-Worth
DOT	Department of Transportation
EADI	Electronic Attitude Director Indicator
EEG	Electroencephalography
EFB	Electronic Flight Bag
EFB_Aft	Avionics Condition showing FIM data on the EFB in the aft position
EFB_Aft+AGD	Avionics Condition showing FIM data on the EFB in the aft position and AGD
EFB_Fore	Avionics Condition showing FIM data on the EFB in the fore position
EICAS	Engine Indicating and Crew Alerting System
FAA	Federal Aviation Administration
FAF	Final Approach Fix
FIM	Flightdeck Interval Management
FMC	Flight Management Computer

FDM/S	Flightdeck Merging/ Flightdeck Merging and Spacing
FMS	Flight Management System
FNL	IMSPiDR indication that algorithm has calculated final approach speed
F/S	Fast/Slow indicator
GIM	Ground-based Interval Management
HITL	Human-In-The-Loop
Hz	Hertz (cycles per second)
IFD	Integration Flight Deck
ILS	Instrument Landing System
IM	Interval Management
IM-S	Interval management-Spacing
IMSACE	Interval Management Systems Avionics Configuration Experiment
IMSPiDR	Interval Management with Spacing to Parallel Dependent Runways Experiment
Integrated	Avionics Condition showing FIM data on the PFD, ND, MCDU, EICAS
KDFW	Airport identifier for Dallas Fort Worth
KLD	Kullback-Leibler Distance metric for scan path comparison
LaRC	NASA Langley Research Center
LNAV	Lateral Navigation
MCDU	Multi-Function Control Display Unit
MCH	Modified Cooper Harper
MCP	Mode Control Panel
MITRE	MITRE company
msec	milliseconds
NASA	National Aeronautics and Space Administration
NASA-TLX	NASA Task Load Index workload instrument
NextGen	Next Generation Air Transportation System
ND	Navigation Display
NNI	Nearest Neighbor Index of visual scan entropy
p	Probability (level of statistical significance)
PF	Pilot Flying
PFD	Primary Flight Display
PM	Pilot Monitoring / Not Flying
POG	Point of Gaze

RMI	Radio Magnetic Indicator
RTA	Required Time of Arrival
RTCA	was Radio Technical Commission for Aeronautics, now simply RTCA
RVT	IMSPiDR display element indicating error when calculating speed
SA	Situation Awareness
SART	Situation Awareness Rating Technique
sec	seconds
SPSS	was Statistical Package for Social Sciences, now simply SPSS
STAR	Standard Terminal Arrival Route
TCAS	Traffic Alert and Collision Avoidance System
TP	Technical Paper
TLX	Task Load Index (NASA-TLX)
TTF	Traffic to Follow
UPS	United Parcel Service
VNAV	Vertical Navigation
VAV	Notification Method with aural for speed deviations only
VSI	Vertical Speed Indicator
VVV	Notification Method with no aural indications

Executive Summary

Flightdeck Interval Management is one of the NextGen operational concepts that FAA is sponsoring to realize requisite National Airspace System (NAS) efficiencies. Interval Management will reduce variability in temporal deviations at a position, and thereby reduce buffers typically applied by controllers – resulting in higher arrival rates, and more efficient operations. Ground software generates a strategic schedule of aircraft pairs. Air Traffic Control (ATC) provides an IM clearance with the IM spacing objective (i.e., the TTF, and at which point to achieve the appropriate spacing from this aircraft) to the IM aircraft. Pilots must dial FIM speeds into the speed window on the Mode Control Panel in a timely manner, and attend to deviations between actual speed and the instantaneous FIM profile speed. Here, the crew is assumed to be operating the aircraft with autothrottles on, with autopilot engaged, and the auto-flight system in Vertical Navigation (VNAV) and Lateral Navigation (LNAV); and is responsible for safely flying the aircraft while maintaining situation awareness of their ability to follow FIM speed commands and to achieve the FIM spacing goal.

The objective of this study is to examine whether three Notification Methods and four Avionics Conditions affect pilots' performance, ratings on constructs associated with performance (workload, situation awareness), or opinions on acceptability. Three Notification Methods (alternate visual and aural alerts that notified pilots to the onset of a speed target, conformance deviation from the required speed profile, and reminded them if they failed to enter the speed within 10 seconds) were examined. These Notification Methods were: VVV (visuals for all three events), VAV (visuals for all three events, plus an aural for speed conformance deviations), and AAA (visual indications and the same aural to indicate all three of these events). Avionics Conditions were defined by the instrumentation (and location) used to present IM information to crews: (1) Integrated (IM information is embedded in extant PFD (Primary Flight Display), ND (Navigation Display), EICAS (Engine Indicating and Crew Alerting System) displays); (2) EFB_Aft (IM information is only supplied in an EFB and mounted in location similar to that for MITRE's UPS work); (3) EFB_Fore (IM information is only supplied in an EFB which is mounted more forward, under the side window), and (4) EFB_Aft plus use of an AGD (the same IM information is supplied in an EFB and on an AGD, both mounted in locations similar to that in MITRE's UPS work). Twelve commercial pilot crews flew descent scenarios (VNAV Speed with the mode control panel (MCP) speed window open until flaps extended, then VNAV Path) in a commercial transport flight simulator with realistic visual scene and communications.

The results of this study serve three practical aims: (1) contribute to the down-select of avionics configuration for future assessment of the ASTAR spacing algorithm at NASA; (2) provide information useful to the FAA Human Factors Division (ANG-C1)'s mission to identify issues pertinent to flight certification of, and flight standards; (3) identify methodological considerations in support of future FIM human-in-the-loop (HITL) investigations.

Significant findings follow, but these should be interpreted in light of data collection and scenario design characteristics as more fully described in the main text:

- Pilots commented positively on the operational concept of IM, and its operational acceptability.
- Pilots' opinions and subjective ratings, as well as some objective measures, favored the Integrated condition; however, it was associated with the most out-of-speed-conformance incidents, primarily when paired with the VVV Method.
- If a retrofit solution is required, it is clear that the EFB_Aft condition was least preferred and resulted in performance decrements – particularly when accompanied with only visual indications.
- In general, more metrics and pilot operational acceptability ratings favor the EFB_Fore condition over the EFB_Aft+AGD condition, however Notification Method interacted with this preference order.
- While Notification Method had some interactive effects across Avionics Conditions, 92% of pilots said an aural indication would be helpful for at least some IM events; and half indicated preference for use in all IM events. Oculometer and other objective results were generally consistent with these subjective data. Pilots who actually experienced aural indications were more likely to recommend them.
- Additional research should address salience of indicators; placement of the EFB and information access in it when concomitant services are provided (especially maps); the algorithm determining conformance deviation alerting to minimize perception of false alarms; and selection of aural indicators to support FIM operations, and evaluation of these in a complete auditory environment.

Future FIM avionics evaluations, should ensure that measures address these objectives:

- support comfortable flight path and speed management;
- permit spare workload capacity to address other aspects of the mission;
- minimize additional workload;
- permit integration of FIM information into an efficient and familiar scan pattern;
- support efficient and error-proof data entry of FIM clearance information;
- provide effectively alerts to exogenously redirect attention as necessary;
- support timely detection of new commanded speeds;
- support extraction of information quickly and unambiguously;
- support timely detection of speed profile conformance deviations;
- provide sufficient notice of profile excursions to avoid deviations and perception of false alarms;
- support effective and timely decisions for IM cancellation and resumption;
- provide well-timed reminders to facilitate speed command entry in a timely manner;
- be consistent with flightdeck design philosophy, cockpit resource management (CRM) and communication conventions;
- not be frustrating, distracting, or physically awkward to use.

1   Introduction

The Federal Aviation Administration (FAA) projects air traffic demand to increase at a rate of 4.9 percent in years 2014 to 2033 (FAA, 2013). Airbus market analyses exceed this, suggesting that commercial transport travel in North America, Western Europe, and Japan, is projected to increase 4% from 2012 to 2031 (Leahy, 2012). Where infrastructure and surrounding populations are established, this increase will demand more efficiency from aviation systems. These efficiencies are most effectively gained by improving terminal airspace operations (Levitt & Weitz, 2011). As such, the FAA is developing NextGen operational concepts and supportive technologies to realize requisite efficiencies (FAA, 2013), and similar development efforts are underway in Europe (Redeborn, 2013). Prior research demonstrated that aircraft arrival rates improve with the precision of spacing between pairs of arriving aircraft (Credeur, 1977). Interval Management (IM) is a flow management concept being developed for near-term implementation to achieve efficiencies in terminal airspace (Levitt & Weitz, 2011). IM operations reduce the variability of temporal deviations at a position, most often applied to arrivals (c.f. Penhallegon Mendolia, Bone, Orrell & Stassen, 2011). By reducing this variability, buffers typically applied by controllers can be reduced, and efficiencies gained in arrival rate. To support continued development of efficient IM operations, this paper first reviews prior research on FIM operations, and then describes this Interval Management Systems Avionics Configuration Experiment (IMSACE), which addressed the efficacy of four display types and three modality methods.

1.1   Improving Terminal Airspace Operations through Interval Management

IM operations are supported by both a ground-based component (GIM) and a flightdeck-based component (FIM). Principally intended for the commercial transport sector, FIM requires both "in" (the ability to receive), and "out" (transmission of) Automatic Dependent Surveillance-Broadcast (ADS-B) equipage on aircraft. While this report focuses on the flightdeck aspect, the GIM component develops the strategic plan for pairing aircraft and establishes spacing objectives. The GIM system derives the content of an IM clearance which the air traffic controller transmits to participating aircraft. It contains:

- the spacing interval to achieve (either in time (seconds) or distance (nm)),
- the target aircraft to follow (TTF)'s identifier, and

- the point along the path at which to achieve the spacing goal (Callentine, Cabrall, Kupfer, Omar & Prevot, 2012; Vu Strybel, Kraut, Bacon, Minakata, Nguyen, Rotterman, Battiste & Johnson, 2010).

When the FIM clearance is received, the flightcrew enters it into an onboard system containing the FIM algorithm. Once the TTF's ADS-B signal is available, the IM aircraft can begin spacing with respect to the TTF. The algorithm uses a speed profile based on the standard terminal arrival route (STAR) in use that conforms to standard speed constraints in the terminal environment (e.g., less than 250 knots under 10,000 feet), and that is adapted to forecast winds. In consideration of these environmental factors, the system displays FIM target speeds that are consistent with the spacing goal, TTF behavior, the intended arrival speed profile, as well as a maximum speed deviation from the intended speed profile. When conducting FIM operations, in accordance with the concept of operations for NASA's technology demonstration efforts (Baxley, Swenson, Robinson, Prevot, Callantine, Scardina & Greene, 2012), the crew is assumed to be operating the aircraft with autothrottles on, with autopilot engaged, and the auto-flight system in Vertical Navigation (VNAV) and Lateral Navigation (LNAV). In FIM operations, the crew is responsible for safely flying the aircraft while maintaining situation awareness of their ability to follow FIM speed commands and to achieve the FIM spacing goal.

1.2   Flightdeck Interval Management Display Research

The RTCA (2011) delineates crew FIM tasks according to initiation, execution, and termination. This study focused on only the execution phase of IM, and so the review of prior work is scoped to this phase. This study does not address the requirements for entering an IM clearance or related information to initiate operations; nor does it address issues associated with termination of FIM operations (e.g., resuming standard operations, suspending IM).

During execution, the flight crew is responsible for: noticing the requirement to enter a new speed; assessing the appropriateness of new speeds; entering new speeds into the flight control system; and monitoring the aircraft's adherence to speeds, progress relative to the spacing goal, and for circumstances which would necessitate termination of IM operations. To support these objectives, avionics must minimally present the crew with:
- target speed information,

- information by which the crew can either monitor progress to the spacing goal or more closely track a speed profile that will achieve this goal, and
- a reference to the TTF that defines the spacing goal.

In addition, this information must be presented such that the onset of new speeds and deviation from the spacing goal / speed profile are sufficiently noticeable to ensure that crews take action "in a timely manner" (specified as an Operational Requirement in RTCA, 2011). A variety of implementations have been studied to provide this information. The following subsections provide an overview of studies that have investigated display concepts supporting FIM.

1.3   NASA Langley

An early study of FIM operations was conducted at NASA Langley (Oseguera-Lohr, Lohr, Abbott, & Eischeid, 2002; Oseguera-Lohr, Lohr, Abbott, Nadler & Eischeid, 2005) in a simulation environment similar to a B757. In this study, pilots were provided with speed target and a fast/slow (F/S) indicator on a modified EADI (Electronic Attitude Director Indicator) (Figure 1). Crews flew in "Speed" mode (speed controlled by dialing the target speed into the Mode Control Panel (MCP) Speed window). When a new speed was presented, a red "speed bug" on the airspeed indicator moved to the target speed displayed in the window, and the F/S indicator reflected the relationship of the current aircraft speed with the target speed. If the current speed was faster than the target speed displayed in the MCP window, the pointer on the F/S indicator moved towards the "F;" if the current speed was slower than the MCP window speed, the pointer moved towards the "S." This modified EADI made use of the F/S indicator to reflect the relationship between the current aircraft speed and the algorithm's intended speed. The EADI also tracked the algorithm speed guidance, and the commanded speed appeared in digital form above the F/S indicator, in green font. When a new target speed was presented, the commanded speed was outlined with a green box and flashed for 5 seconds. The displayed readout, the pointer on the F/S indicator, and the bug on the airspeed indicator all followed the commanded speed from the FIM algorithm (Osequera-Lohr et al., 2002).
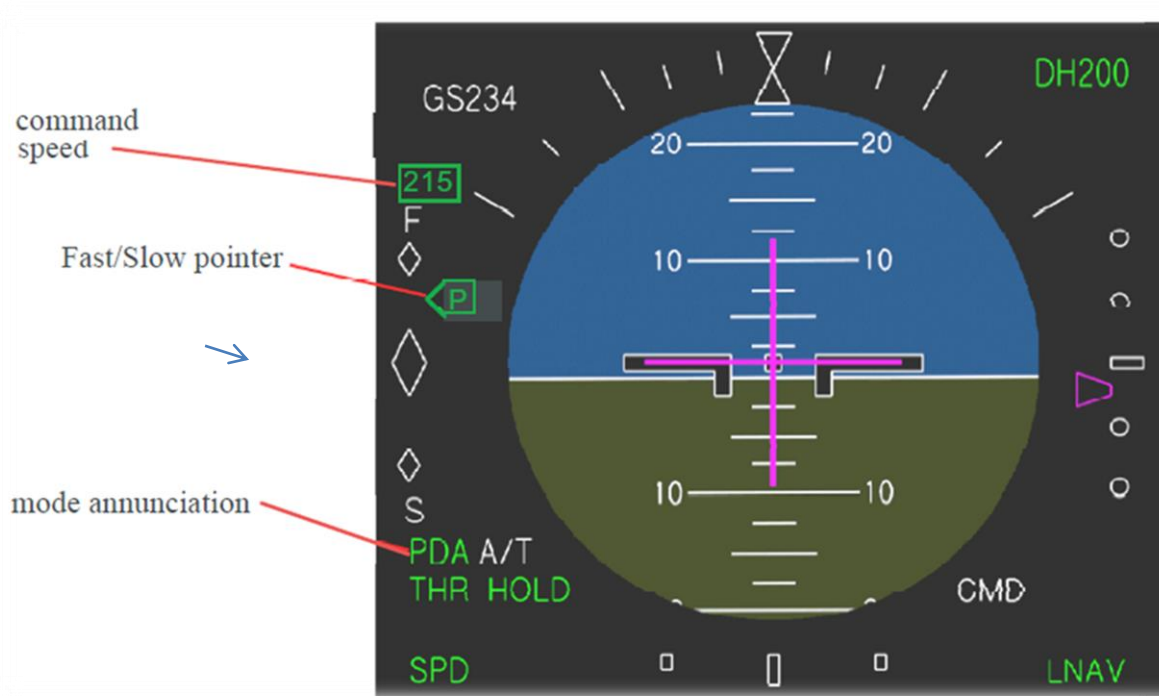
Figure 1. EADI in Osequera-Lohr et al. (2005)

The navigational display also was modified (Figure 2) to include: 1) a data block that displayed currently entered FIM data and TTF range, 2) a spacing position indicator, and 3) TTF highlighting and position history dots. The position indicator showed where the IM aircraft should be in order to achieve the spacing goal, and consisted of a green line perpendicular to the IM aircraft's ground track, with an inverted "V" at the line's midpoint. When properly spaced, a white triangular IM aircraft symbol was displayed. If the spacing position indicator was behind the apex of the IM aircraft symbol, the IM aircraft was ahead of where it should be (actual spacing interval was less than the targeted interval). Conversely, if the spacing position indicator was ahead of the IM aircraft symbol, then the IM aircraft was behind where it should be (actual spacing interval was greater than the targeted interval). This indicator provided a visual reference of the IM aircraft's position relative to the desired spacing interval. In this concept, the Multi-Function Control Display Unit (MCDU) was adapted to enable entry of the IM clearance information, final approach speeds of IM aircraft and TTF, and winds.

4

Figure 2. ND in Osequera-Lohr et al. (2005)

A human-in-the-loop (HITL) investigation of this implementation showed no significant differences in reported workload for a FIM approach versus standard flight operations. The results also showed that heads-down time with the FIM was acceptable and there was minimal impact on the pilots' nominal scan patterns. In general, the evaluation crews' commented positively on the FIM displays. High ratings were given for the amount of symbology, the effectiveness of the commanded speed to communicate fast/slow, for detection of a new speed, and for highlighting of the TTF. Ratings on the utility of the spacing indicator ranged significantly, and some pilots indicated they did not use it. The F/S indicator on the EADI was rated only moderately effective. While the blinking commanded speed indicator was highly rated, some pilots indicated that they would have preferred that it blink for longer, or until the speed was changed. In a subsequent flight test with this implementation, pilot comments were similarly positive with regard to overall acceptability, heads-down time, and confidence in using the FIM tool and displays (Lohr, Oseguera-Lohr, Abbott & Capron, 2003). The algorithmic underpinning this evaluation began development in the early 1980's (Abbott & Moen, 1981), and has evolved to support current day FIM concepts (Abbott, 2013).

The Interval Management with Spacing to Parallel Dependent Runways (IMSPiDR) experiment was conducted in 2011 to evaluate new display concepts andhe Airborne Spacing for Terminal Area Routes spacing algorithm (ASTAR) that could calculate trajectory and spacing calculations for two aircraft on parallel, offset runways and was designed to accommodate late (after established on final) changes to aircraft separation criteria (Baxley, Swenson, Prevot & Callantine, 2012). In this study, several simulation environments were used concurrently to assess the feasibility and behavioral implications of this algorithm in multi-user scenarios. New display concepts for the Primary Flight Display (PFD) and Navigational Display (ND), and MCDU were developed for that environment and are depicted in Figure 3 (Swieringa, Murdoch, Baxley & Hubbs, 2011).

Figure 3. IMSPiDR Displays (Swieringa et al., 2011)

The PFD used in IMSPiDR (Swieringa et al., 2011), retained all normal display elements (airspeed, altitude, etc.), and added three elements to support airborne spacing operations (Figure 3, upper left). Clockwise from the bottom left of the PFD, a green speed bug displayed the commanded speed from the spacing algorithm (the magenta speed bug corresponded to the commanded speed in the Flight Management Computer (FMC)) and a three digit green number displayed the commanded end speed from the spacing algorithm (the magenta three digit number immediately below it corresponded to the FMC commanded speed). To the right of the commanded end speed, letters identified which of the four modes the ASTAR algorithm was in: "RTA" (for "Required Time of Arrival") indicated the algorithm was using the RTA control method at the runway threshold to calculate the commanded end speed, "IM" indicated the algorithm had valid ADS-B data on one or both of the lead aircraft and was calculating a speed to achieve

the assigned spacing interval, "FNL" indicated the algorithm was displaying the FMC calculated final approach speed, and "RVT" indicated an algorithmic or operational error. On the ND, lead aircraft (as there may be two) were denoted by a white diamond, and an aircraft call sign was included in each lead aircraft's expanded data tag (Figure 3, upper right). During IM operations, the lead aircraft that was currently controlling the spacing had one of its double chevrons/diamonds changed from white to green. The bottom row of displays in Figure 3 show new pages added to the Flight Management System's (FMS) MCDU. The first page displayed summarized information about the lead aircraft, including spacing error, runway, final approach speed, spacing error, and aircraft call sign. The first page was also used to activate the FIM procedure and, when necessary, to terminate the operation. Two additional IM pages contained supplementary information regarding each lead aircraft, including their IM-S goals, achieve by points, terminate-at points, and routes. Finally, the Engine Indicating and Crew Alerting System (EICAS) display provided messages that indicated conditions with respect to IM operations.

The HITL study using these displays investigated the effect of different control methods (RTA and IM) and different sources of error (a baseline with no error, winds, and a runway threshold offset). Results indicated that pilots found the operations acceptable in terms of workload distribution as well as crew coordination; but, pilots were not as positive regarding the extent to which IM operations could be integrated with current procedures. In addition, some workload level ratings were unacceptably high. Further, pilots appeared to intentionally accept large speed conformance errors to avoid the fuel implications and passenger discomfort that would attend minimizing these. Evaluation of display features indicated that the PFD elements were all 'very useful' with the exception of the green box around the commanded speed (which was used to indicate the occurrence of a new commanded speed). Results indicated that this indication was too subtle for the vast majority of participants. Sixty-seven percent wanted an aural alert to notify of speed changes, and 54% wanted the box on commanded speed to flash. The spacing trend indicator on the ND was rated as one of the least relevant design elements. Pilots were also shown a "conformance box" (a box around the ownship that conveys the spatial extent to which ASTAR can still accomplish the spacing goal) on the ND and asked their impressions of this display element. Crews reported that the displays were acceptable, but improvements were desired.

In addition to the present study, additional research concepts are under investigation at NASA Langley. NASA's Air Traffic Management Technology Demonstration-1 (ATD-1) is the most

recent, full implementation of an IM concept (Prevot, Baxley, Callantine, Johnson, Quon, Robinson, & Swenson, 2012; Baxley, Swenson, Robinson, Prevot, Callantine, Scardina & Greene, 2012). Towards this end, several studies have evaluated IM operational procedures using a mid-term implementation involving the use of an Electronic Flight Bag (EFB) and an auxiliary ADS-B guidance display (AGD), and have found workload acceptable (Shay, Swieringa & Baxley, 2012; Murdoch, Wilson, Hubbs & Smail, 2013; Wilson, Murdoch, Hubbs & Swieringa, 2013). Novel display concepts are also evolving. Most recently, a novel auxiliary display has been developed to present FIM information in the forward field of view, in conjunction with use of the EFB (Swieringa, 2013) to replace the AGD. This display includes a fast/slow indicator, the commanded end speed mode, TTF identifier, system messages, and a progress indicator (indicating if the ownship is early or late relative to the spacing goal).

1.4   Eurocontrol

Hébraud, Hoffman, Pene, Rognin & Zeghal (2004) summarize a series of three HITL studies investigating the Eurocontrol CoSpace algorithm and displays. The earliest study of this concept had no additional guidance beyond a basic Cockpit Display of Traffic Information (CDTI) display and resulted in higher workload ratings (Bone, Penhallegon & Stassen, 2008a). The second study augmented the ND with the spacing scale and a suggested speed indicator. While the suggested speed seemed to be an improvement, pilots suggested that the closure rate information element was not useful. The third study further augmented the displays (Figure 4) with additions including: representation of the TTF position and heading with a data tag specifying relative altitude and vertical trend; a reference line that linked the IM aircraft to the TTF and indicated the point where the trajectories would merge; flight mode; the location at which the spacing objective should be accomplished; the current and required spacing with trend information (30 seconds look ahead), closure rate, and tolerance margin; and the suggested airspeed to accomplish the spacing goal. When the speed tolerance exceeded 7 knots, the suggested speed would blink. Speeds were color coded to indicate feasibility; green if feasible, amber if flaps are required, and red if outside aircraft performance. This final concept also included advisories and caution indications on the ND to alert the crew to stabilize speed, accelerate/decelerate, direct-to advisories and 'unable to continue' advisories. In this final evaluation, returning pilots rated these augmentations as improvements, and new participants were complimentary. Spacing trend information and the blinking suggested speed indicator

received positive comments, although some pilots indicated that the suggested speed indicator should be more salient. Closure rate and color coded cautions were not as well received.
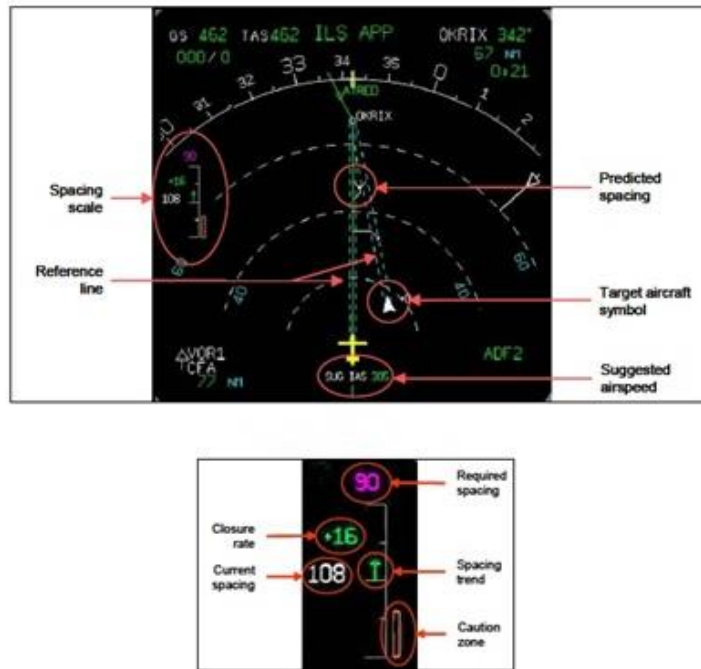
Figure 4. Eurocontrol Study 3 Display Content (Hébraud et al., 2004)

1.5 MITRE

MITRE tested an implementation planned for limited United Parcel Service (UPS) installation. This implementation included a Traffic Collision Avoidance System (TCAS) display (which provided TCAS target information without any traffic or resolution alerts), an EFB with cockpit display of traffic information (CDTI) mounted below the windows and aft of the tiller, and an AGD mounted under the glareshield (Figure 5). The CDTI on the EFB provided typical information found on a ND as well as ADS-B traffic information; and when in IM operations, showed a fast/slow speed bug and the current and commanded Mach speeds (Figure 6). The touchscreen enabled display allowed selection of targets which would present additional information about a selected aircraft (flight identification/call sign, range, ground speed, aircraft category). The AGD presented commanded speed, and the TTF's range and TTF's differential ground speed.

Results of piloted simulations with this configuration resulted in mixed ratings regarding workload, with some reports indicating higher workload than in a baseline condition, but still at an acceptable level (Bone et al., 2008a). Pilots reported appropriate situation awareness to understand spacing and progress towards the spacing goal, as well as their ability to detect evolving situations. Bone et al (2008a) reported increased heads-down time over the baseline condition, but considered the extent of this not concerning. While all pilots found the location of the AGD acceptable and easy to include in scan, the location of the EFB was not universally accepted. Pilots reported using the EFB CDTI more than the AGD. Further, there was some confusion regarding the presentation of ground speed on the EFB. Bone et al. (2008a) attribute this to the novelty of the CDTI display, and the need for a richer information source for non-normal conditions than was provided by the AGD. All but one pilot said this configuration was acceptable for conducting FIM operations. While overall heads down time appeared to be acceptable to these investigators, they concluded that the time spent viewing the CDTI was not acceptable, and suggested further investigation of this problem. Pilots were asked about whether an aural alert would be useful to indicate new commanded speeds. In an earlier study (Bone, Helleberg, Domino & Johnson, 2003), Bone reported that pilots requested aural alerts with a CDTI after having experienced them.

In a subsequent study (Bone, Penhallegon & Stassen, 2008b), this display configuration was extended to investigate continuous descent arrivals. The display for this experiment (Figure 7) was similar to the prior but included a spacing deviation indicator which showed the IM aircraft's position relative to the required position to achieve or maintain the desired spacing goal. In this implementation, the AGD was redesigned to present only the commanded speed and the range of the TTF. When a new commanded speed occurred, the speed value on the AGD flashed for 5 seconds. If the new commanded speed was more than 10 knots different from the current speed, the "CMD" text next to the commanded speed flashed for 10 seconds. If the deviation between selected and commanded was less than 20 knots, flashing stopped automatically when the commanded speed was entered. If the difference between the entered speed and the commanded speed was greater than 20 knots, the flashing persisted. If the pilot wished to extinguish the flashing in any circumstance, this could be accomplished by pressing a button on the device. The avionics in this study were in the same locations as in the Bone et al. (2008a) study. Results of this study showed pilots considered the FIM test scenarios using this equipment to be relatively low workload. Most participants considered situation awareness to

be acceptable; and, while they generally agreed that FIM introduced more heads-down time, most reported that this was not unacceptable. Most participants agreed that this configuration was acceptable to conduct FIM, and all agreed that the displays provided the required information for the operation, including an understanding of target spacing and progress towards the spacing goal. In particular, participants suggested that the selectable target data block and symbol were very useful, whereas the fast/slow bug and commanded speed were only moderately useful. The progress icon and the vector lines received mixed reviews. Generally, pilots liked the AGD and found the commanded speed on this device very useful. Pilots rated the deviation information as less useful. Only one pilot suggested that an aural indication accompany a new speed command. Pilots suggested that an aural alert would be appropriate for indicating that a commanded speed had not been dialed in after a period of time. For those who reported a preference, it was suggested that a tone would be appropriate. It was noted by some pilots that overuse of aural alerts could be annoying.

Figure 5. Location of Displays for MITRE Study (Bone et al., 2008).
(The thinner circle indicates location of EFB, the thicker indicates location of the AGD)

Figure 6. MITRE Display Content for FDMS-2 (Bone et al., 2008a)

Figure 7. MITRE Display Content for FDMS-3 (Bone et al., 2008b).

Penhellagon, et al., (2011) conducted a HITL study in which EFBs (with CDTI) are described to be at the pilot's eleven o'clock position, and at the first officer's one o'clock position. This study also used an AGD positioned under the glareshield as before. Based on review of prior literature, the authors concluded that the information requirements for FIM operations include: TTF identification, the assigned spacing goal, IM speed, and when appropriate a merge point. This implementation preserved the ability to select target information as above. The AGD included the commanded speed, TTF identification, and the current spacing goal expressed as In-Trail Time. When a new speed command occurred, the speed on the AGD displayed a green outlined box for 10 seconds. Generally, pilots found flying FIM departure scenarios with this equipment to be acceptable, and found the information sufficient to detect two off-nominal conditions (one in which speeds outside the operating limits of the aircraft were given, forcing

13

crew to terminate IM; the other suspended-and-resumed IM operations, as initiated by ATC). As in past studies, evaluation pilots reported an increase in situation awareness. While this study cites its results and others as underscoring the need for the CDTI, pilots in the pilot flying role suggested they would focus only on the AGD, whereas those in the pilot monitoring role desired the CDTI information. On the AGD, the commanded speed was the most important element, followed closely by the spacing goal information. Penhallagon et al (2011) concluded that salience of the visual indications was high, based on high speed conformance, but some pilots suggested an aural indication would be helpful to accompany new speed commands. On the CDTI, the differential ground speed information was most desired (more than TTF range), and this seemed to be used by some pilots' to assess feasibility of speed changes.

1.6 NASA Ames

Prevot, Callantine, Kopardekar, Smith, Palmer & Battiste (2004) described the coordinated concept of trajectory-oriented operations, and presented a CDTI with guidance regarding spacing. In this display concept, the box and coloration indicated if the ownship is behind (white), in (green) or ahead (amber) of the range of positions that will achieve correct spacing. This display element was integrated into a 3d-"volumetric" NASA CDTI in a test focusing on procedures and system behavior (Prevot, Callantine, Homola, Lee, Mercer, Palmer & Smith, 2007), and a spacing panel was provided that depicted the information derived from the IM clearance (Figure 8). This study focused on system performance and controller workload. No results were provided to characterize pilots' use of this display set.

1.7 UPS Field Evaluations

Aforementioned work by MITRE was used to develop IM protocol and equippage specifications for the UPS field evaluations. Moertl and Pollack (2011) describe a series of flight tests
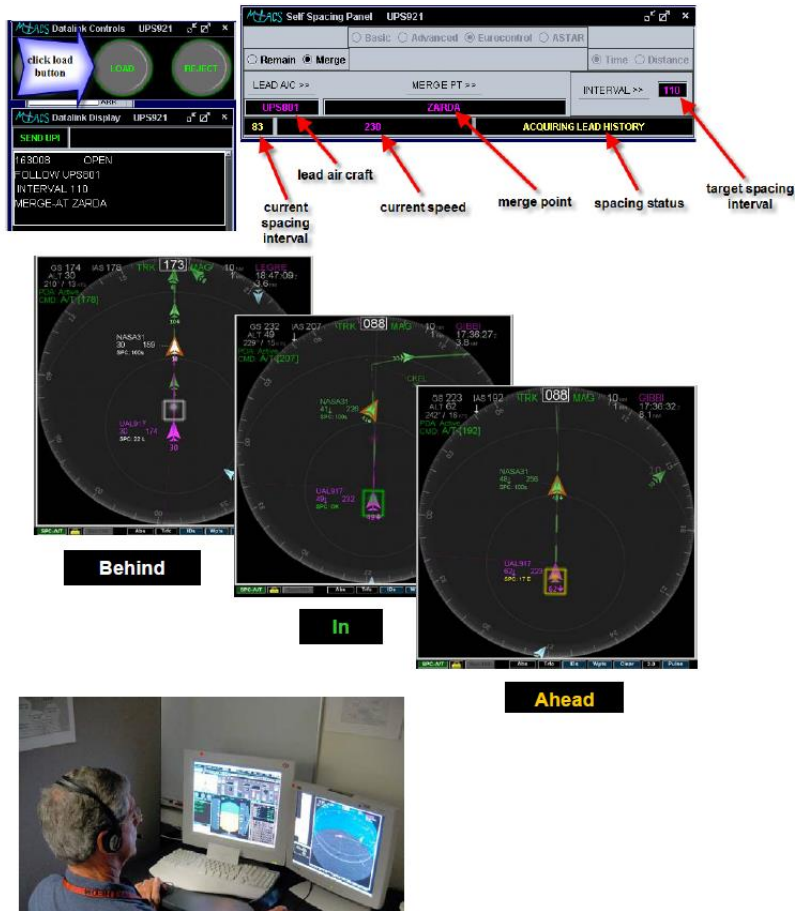
Figure 8. NASA Ames CDTI for IM Spacing (Prevot et al., 2007).

that involved the UPS air carrier to assess various IM software implementations. The displays are described (not depicted) to include flight identification, aircraft type, departure airport, predicted spacing, predicted fix crossing times, flight plan information, and indicated airspeeds. Pilots were notified of speed advisories via visual cues only. Moertl and Pollack's reported results were limited to system performance effects, but mention is made here for completeness.

Most of the previously mentioned research efforts have provided FIM supportive information in extant PFD, ND, EICAS displays; integrating this information into the existing flightdeck. The MITRE (UPS) work used other displays – an EFB mounted aft of the tiller to the side, and an AGD mounted under the MCP, and the Langley IMNOVA used an AGD in this same position. While an integrated solution may seem most effective, it also requires a significant investment in development and certification. As such, retrofit solutions have been suggested for the mid-term

15

(2025) implementation of IM. Those studied thus far include the use of an EFB (or auxiliary display) and an AGD. This study, IMSACE, was conducted to support down-selection of avionics for NASA's Advanced Technology Demonstration (ATD) of FIM operations. As such, display concepts presented in the Avionics Conditions were inherited from prior FIM studies at NASA Langley (i.e., IMSPiDR and derivatives), and used EFB display formats developed for this ATD.

Prior research has discussed the utility of adding aural alerts to support FIM operations (Swieringa et al., 2011; Bone, et al., 2008a; Penhellagon, et al., 2011), but this notification method has not been tested. Notification Methods, were developed for the present study (IMSACE) to test the effectiveness and necessity of aural indications for FIM events.

2 Objectives

This study investigates four Avionics Conditions:

- EFB_Aft (IM information is only supplied in an EFB and mounted in location similar to that for MITRE's UPS work);
- EFB_Fore (IM information is only supplied in an EFB which is mounted more forward, under the side window);
- EFB_Aft plus use of an AGD (the same IM information is supplied in an EFB and on an AGD, both mounted in locations similar to that in MITRE's UPS work); and
- Integrated (IM information is embedded in extant PFD, ND, EICAS, and MCDU displays).

The first three of these are considered retrofit options, whereas the Integrated condition requires redesign of primary flight instruments. These conditions are depicted and described in more detail in the Methods section.

This study also tests the use of visuals and aurals for notifying pilots to the onset of a speed target, conformance deviation from the required speed profile, and for a reminder when a speed target is not yet entered. Notifications always included a visual indication (V); but in some cases, the visual was accompanied by an an aural indication (A). Three Notification Methods were used, and are hereafter referred to as:

- VVV (speed onset=V, deviation=V, reminder=V);
- VAV (speed onset=V, deviation=A, reminder=V)); and
- AAA (speed onset=A, deviation=A, reminder=A).

More details on the construction of the Notification Methods are presented in the Methods section.

The main objective of this study is to ascertain the relative advantages and disadvantages of these Avionics Conditions and Notification Methods for supporting FIM operations, in the context of arrivals for commercial aircraft crews.

Results of this study are in service of three practical aims: (1) to contribute to the iterative design process and down-select of avionics configuration for future assessment of the ASTAR

spacing algorithm at NASA; (2) to provide information useful to the FAA in their mission to identify issues pertinent to flight certification of, and flight standards for FIM operations; (3) to identify methodological issues that may be helpful to future FIM human-in-the-loop (HITL) investigations.

3 Methodology

3.1 Evaluation Pilots

Twelve crews, each consisting of one Captain and one First Officer from the same airline, participated in the experiment. Pilots were type rated in the same class as the simulated aircraft. One evaluation pilot failed to provide experience data. The remaining evaluation pilots had between 19 and 48 years of flying, with the mean and median of these data both 30 years. These pilots had between 400 and 23,600 commercial flight hours, with the mean and median of these data being 7,115 and 4,670 hours, respectively.

3.2 Independent Variables

The Avionics Configurations tested were defined by an Avionics Condition (display devices and locations) and a Notification Method (whether indications of events were presented only visually, or were augmented with aural indications). Avionics Configurations are decribed in terms of levels below, with subsequent sections presenting more detail and pictures of these.

Each crew evaluated four Avionics Conditions: (1) Integrated – in which the FIM target speed was presented in the upper left corner of the PFD and speed profile deviation information was implicitly indicated as the deviation between current speed and an instantaneous speed profile bug on the PFD speed tape. In this Avionics Condition, clearance information would have been entered, and could be referenced, on the FIM page on the MCDU. This page also gave a digital readout of speed profile deviation. Significant deviations from the speed profile triggered a message on the EICAS system. (2) EFB_Aft – in which an aft-mounted EFB was used as the device housing the FIM algorithm and presented all the relevant information for the operation, including speed targets, speed deviation information, and all elements of the IM clearance. Significant deviations from the speed profile triggered messages on the same EFB display. (3) EFB_Fore – in which the same information was presented as in the EFB_Aft condition, but the display was mounted in a more forward location, just under the outboard window. (4) EFB_Aft+AGD – in which the EFB_Aft condition was augmented with the ADS-B Guidance Display (AGD). The AGD repeats the same FIM target speed and speed deviation information on the EFB.

Crews received notifications when conditions required their attention, *i.e.*, when a new FIM target speed occurred (*target speed onset*), if the current aircraft speed significantly deviated from the FIM target speed (*conformance deviation*), and if they failed to enter a new FIM target speed within a reasonable time period (*reminder*). A conformance deviation indicator was provided when all these conditions were true: the aircraft current speed was more than seven knots different from the instantaneous speed on the FIM speed profile for more than twelve seconds, the speed changed more than five seconds ago, and aircraft current speed was not converging to the FIM target speed. A reminder was provided if the crew did not dial in the correct FIM target speed within 10 seconds. If the speed was still not dialed in, the reminder indication was repeated at most two more times at 10 second intervals.

This study evaluated three notification methods defined by the modality (V for visual, A for aural) associated with the triplet of implementations: *target speed onset*, *conformance deviation*, and *reminder* events. The VVV method provided only visual (V) cues for all three events. The AAA method augmented these visual indications with an aural (A) tone, again for all three events. The VAV method included visual indications for all three events, and presented the aural indication only if pilots significantly deviated from the speed profile (conformance deviation). In sum, these are defined as follows:

- VVV (speed onset=V, deviation=V, reminder=V);
- VAV (speed onset=V, deviation=A, reminder=V)); and
- AAA (speed onset=A, deviation=A, reminder=A).

Each crew member had the opportunity to fly an arrival and approach with each of the Avionics Conditions twice, once as pilot flying (PF) and once as pilot monitoring (PM).

3.3   Experimental Design

The experiment had three factors:

(1) Avionics Condition (with four levels: Integrated, EFB_Aft, EFB_Fore, EFB_Aft+AGD) – (a within-crew variable);

(2) Notification Method (with three levels: VVV, AAA, VAV) - (a between-crew variable);

(3) For flight performance measures, a term was added to investigate effects of seat position of the PF v. PM role (a within-crew variable), and for post-run questionnaire items, models instead included a term to investigate effects of pilot role (PF, PM).

The order of Avionics Conditions was counterbalanced over crews, and the assignment of scenarios to Avionics Conditions was arranged so that the pilot flying never experienced the same runway approach twice in a row. Pilots switched roles as PF and PM every other run, but did not change seats to do so. Notification Methods were indexed over days to minimize experimental; bias.

3.4   Facility & Apparatus

Simulation Facility

The Integration Flight Deck (IFD) Simulator is similar to a Boeing 757-200 aircraft cockpit, with a similar aircraft dynamics mathematical model (Figure 9). The flightdeck includes standard instruments representative of a line operations Boeing 757-200 aircraft, including: a pilot-side heads-up display (not used in this study), two MCDUs, a column/wheel control loader system for pitch/roll control, and a set of control loaded rudder pedals. It has a panoramic, collimated external scene Evans and Sutherland 4530 display, providing 200 degrees horizontal field-of-view and 40 degrees vertical field-of-view (Bailey et al., 2006). For this experiment, the IFD was used in a fixed-base configuration.

Figure 9. NASA Langley's Integration Flightdeck Simulator

The IFD is equipped with two Smarteye Oculometer systems, comprised of twelve small cameras and eight flashes that operate in the infrared spectrum (outside perceptible range). The Smarteye system reported head and eye position, quality metrics associated with head position and gaze direction, and the AOI that the reported POG intersects for both pilots.

FIM Avionics

The set of displays that were used in this study to display IM information included, PFD, ND, EICAS, and MCDUs; as well as two EFB, and AGD designed specifically to support IM. Figure 10 shows the EFBs in the fore position, and depicts the approximate location when placed in the aft position (which is approximately as implemented by MITRE, shown in Figure 5). Not all displays were used for all experimental conditions. The sets of display conditions studied were described in the Independent Variables section. For the integrated condition, the display formats used for this study are similar to that used in the IMSPiDR study; however, a box that flashes was implemented around newly received commanded speeds. Parallel runway symbology was used, and did not use a conformance box. Only one CDU page was used (the first in Figure 3), and the EICAS message set was reduced to "IM SPACING" as a mode indicator. The other Avionics Conditons in this study use equipment similar to that used in the MITRE series of studies; an AGD and EFB. The EFBs also contained arrival and approach charts. Pilots had paper copies of this information as well, but pilots generally did not use this, and referred to this information onthe EFBs. Appendix A depicts the implementation of IM information on each display, and the position of displays.

For the conditions using EFBs, new speeds were annunciated visually by the speed value turning reverse video until the pilot dialed-in the speed. If the pilot did not enter the speed in the MCP within 10 seconds, a reminder was provided in the form of a flashing speed indicator. Deviations from speed profile were displayed by text showing the number of knots the pilot was fast or slow, where (slow deviations were denoted as negative values). For the EFB condition, if the pilot was out of conformance with the speed profile, the speed would flash. The EFB_Aft+AGD condition was the same as the EFB condition with the addition of this information on the AGD: illumination of a small white light and a change in the value in the speed text box;

2) flashing speed reminders, and 3) speed conformance deviations blink the fast/slow value. In the Integrated condition, commanded speeds were shown in the upper left corner of the PFD. When a new speed command occurred, a green box was drawn around the speed value, and if a reminder was required, this box would blink. In this condition, conformance to the speed profile was indicated as a "bug" (a small arrow pointer) to the commanded speed on the speed tape and as a fast/slow indication in text on the MCDU IM page. The PFD speedtape bug presented a continuous indication of deviation from the intended profiles speed. In addition, if a conformance deviation (meeting the logic associated with tones above) occurred, the message ("IM SPD DEV") was presented on the upper EICAS display.

In some cases, the visual indications were augmented with an aural indication. The tone used was designed in accord with AC 25.1322-1 guidelines (FAA, 2010), and prior literature regarding urgency and annoyance (Gonzalez, Lewis, Roberts, Pratt & Baldwin, 2012). The aural indication (Appendix B) consisted of two identical tones - each with peak energy at 300 Hz (recommended to maximize urgency, with minimal annoyance), 1050 msec in duration, and separated by a silent interval of 450 msec. Sound pressure level in the simulator during tone presentation was approximately 70 dBa.
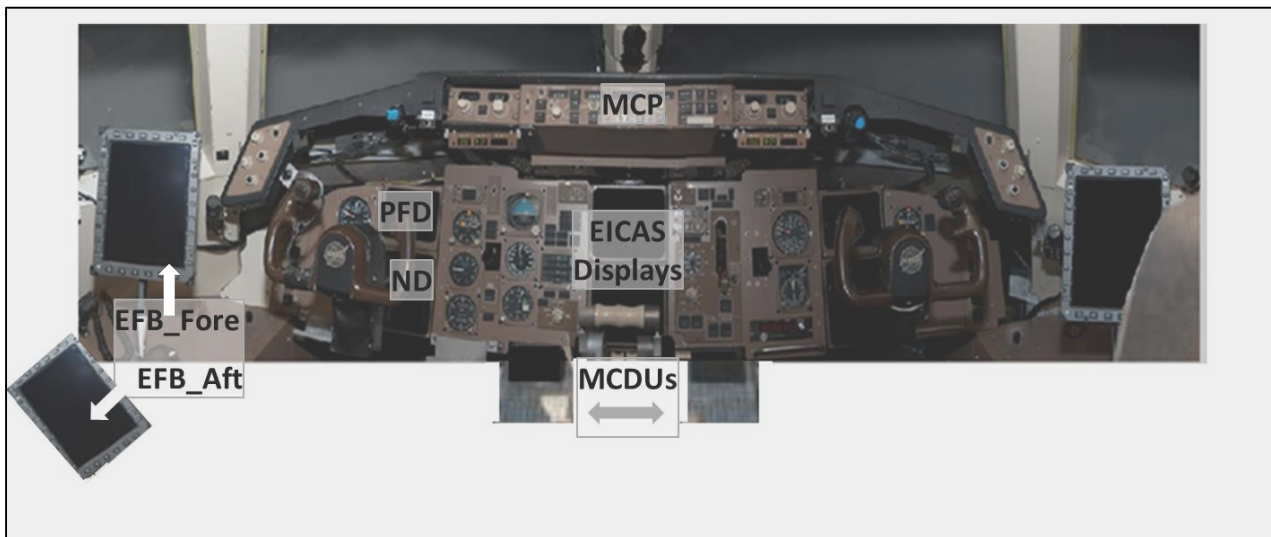
Figure 10. Interval Management Avionics

(Note: All displays are available to both sides of the flightdeck)

3.5 Environment & Scenarios

Scenarios required crews to fly an arrival from an altitude of approximately 25,000 feet to land at Dallas-Fort Worth International Airport (FAA identifier: KDFW). This study used modifications of the Bonham Five, Cedar Creek Six, Bowie One, and Glen Rose Nine arrivals to runways 17C and 18R. For this study, these arrivals were connected to the approach to permit IM operations to the Final Approach Fix (FAF), with a final speed provided at approximately 1000' (AGL) permitting a stabilized approach. Appendix C provides the details on the environment. These arrivals were flown in the context of moderate to heavy traffic as recorded during actual operations at DFW, with both IM-equipped and non-participating commercial aircraft.

Scenarios began in level flight prior to top of descent, with VNAV Path autoflight mode engaged. Crews were given their IM clearance verbally and in written form. Crews were not required to enter the IM clearance information, as this would normally have been done at an earlier point in the arrival. The aircraft was flown in an unconstrained vertical path descent from the Top of Descent until reaching the first altitude constraint at 11,000 feet. The aircraft operated in VNAV Speed with the MCP speed window open until the flaps were extended, and the autoflight mode reverted to VNAV Path. Crews received their first FIM speed target within the first three minutes of the scenario. When the FIM system provided a speed change, the pilots implemented the new speed by setting it into the MCP speed window. The last speed target given was the reference speed for flaps at 30 plus 5 knots to enable stabilized approach by 1000 feet above ground level. Crews typically flew the Instrument Landing System (ILS) glideslope to the runway threshold, and were permitted to continue to rollout. However, in some instances it was necessary to accelerate the test schedule by concluding a data run early, but never before the aircraft was configured for a stabilized approach. Crews were instructed to: (1) fly as they typically would, as though they had passengers; (2) respond to speed targets in a timely manner; (3) maintain speed conformance within 7 knots (using IM avionics), and; (4) remain within 400 feet of the VNAV path (using the vertical path deviation indicator on the PFD, and/or information in the MCDU).

Confederate Air Traffic Controllers provided realistic communications to both the simulator and to the simulated aircraft in the environment. These communications were consistent with current operations and with IM phraseology from a draft version of the ATD-1 Concept of Operations (Baxley et al., 2013). Prerecorded Automatic Terminal Information Service (ATIS) messages were available on the appropriate frequency. During the scenarios, controllers indicated when a change in ATIS occurred, and a subsequent version was available with no substantially challenging conditions. Wind speeds and directions were constant and not challenging.

3.6 Experimental Protocol

Crews arrived at NASA LaRC in the morning and were shown the IFD simulator, the IM equipment, and the oculometer apparatus in the context of a brief discussion of the scenarios. Following this discussion, crews were given information about the experiment, provided informed consent, and received detailed classroom training. The classroom training included: a description of the data collection instruments; the experimental scenario and environment; the conceptual method used by the FIM algorithm to generate target speeds; how arrivals would be flown in accordance with IM avionics and the information and indications that would be provided to support these operations; and guidance on performance goals.

After classroom training, the pilot trainer reviewed all instrumentation and scenario requirements with the evaluation pilots in the IFD simulator. Two 30 minute training runs were flown for familiarization. For both training runs, the pilots switched roles at an intermediate point so that each pilot could conduct IM operations in the more intense environment at lower altitudes and when closer to the airport.

After the training run, crews conducted nine data collection runs. The first eight runs consisted of two runs per Avionics Condition, wherein the pilots switched roles between PF and PM. The final run was not analyzed in this report. Each run was approximately 20 minutes long, with an additional 5-10 minutes required to complete a post-run questionnaire. At the end of the day, each crew member completed a post-experiment questionnaire and participated in a debriefing discussion. Total time for an evaluation crew was nominally nine hours.

3.7   Data Collection Instruments & Dependent Variables

Four forms of data were collected in this study: (1) simulation data, (2) post-run questionnaire data, (3) post-experiment questionnaire data, and (4) oculometer data.

Simulation Data

Simulation data (aircraft performance and position, control inputs, IM events) was collected at 5Hz during each run. Simulation data contained timestamped parameters that were used to assess vertical deviation from the VNAV path, speed deviation from the intended IM profile, and use of speedbrakes and throttle to manage vertical path during speed changes. Simulation data also contained data to represent the amount of time the aircraft was in the state defined by the conformance deviation indications, the number of out-of-speed-conformance (speed deviations) that occurred, as well as the number of IM reminders that occurred. Simulation data were also used to determine crew response time; i.e., time from the occurrence of a speed target until they began dialing the speed into the MCP.

Post-Run Questionnaire Data

A post-run questionnaire (Appendix D) was completed immediately following each run to assess workload components pertaining to overall demand level (Casali & Wierwille, 1983), and specifically mental and temporal demand, performance, effort, and frustration (Hart & Staveland, 1988) situation awareness (Taylor, Selcon & Swinden, 1995) and various questions pertaining to the acceptability of the IM information received and scenario demands (7-point scales with "Completely Agree" and "Completely Disagree" anchor cues).

Post-Experiment Questionnaire Data

Appendix E contains questions from the post-experiment questionnaire pertaining to this report. These include questions that asked the pilots to indicate the degree of preference for an Avionics Condition, given each possible pair of Avionics Conditions, to rate the operational acceptability of Avionics Conditions; and to assess the utility of aural indications, use of speedbrakes, appropriateness of thresholds for IM reminder and conformance deviation indicators, operational acceptability with regard to cockpit procedures and cockpit resource management, and effect on workload and situation awareness. This questionnaire also asked crews to assess the effectiveness of the training they received and the realism of the simulation and scenarios. Exploratory questions regarding use of FMC-coupled speed targets, use of the

EFB for IM and charts, and placement of IM information were included, but are not analyzed in this report.

Oculometer Data

Oculometer data were collected by the coupled Smarteye oculometer systems (Latorella, Ellis, Lynn, Frasca, Burdette, Feigh & Douglass, 2010) operating at 60Hz. Point-of-gaze (POG) data is derived from eye position and head position data, and is futher translated into metrics by which to compare effects of Avionics Conditions and Notification Methods. These include: visual sampling of FIM displays, Heads-Up sampling, time to notice new speed targets (for each pilot, and the minimum for the crew), workload-related measures (based on dwell time and regularity of scan), and a measure of the similarity between pilot's scan paths. Measures associated with sampling of the FIM displays and Head-Up sampling were conducted on the majority of data in a run (99 seconds into the run until the end of the run and normalized for run length). Analysis of other measures used event-locked periods which were anchored on the occurrence of a new FIM speed, and included data for the following 19 seconds. This duration was chosen to ensure that, for all cases, these windows did not overlap.

Oculometer data was recorded in logfiles at 60Hz, which corresponds to a nominal frame rate of about 17 msec. Oculometer data was also sent to simulation files, which were recorded at 5Hz. This experiment resulted in 192 oculometer logfiles (12 crews x 2 pilots/crew x 8 runs/crew), 87.5% of which (i.e, 168 logfiles) were suitable for further analysis. The majority of missing data pertained to the first speed target, after pruning these from consideration in all datafiles, only seven logfiles (approximately 3.6%) were affected by significant data loss.

In addition to incomplete data files, recorded data may be of questionable quality due to individual eye and eyelid characteristics and postures. Smarteye's software reports a head and gaze quality value for each reported point of gaze (POG). This gaze-quality metric is defined by the system's confidence in head and eye position assessment, normalized over the data previously acquired in that session. The gaze direction quality value ranges from 0.0 to 1.0; where 0.0 corresponds to the 1st percentile of all quality values experienced to that point, and 1.0 corresponds to the 99th percentile. As such, this value is individual-dependent and only useful as a general guide to the degree to which the oculometer has sufficient information upon which to base a POG determination. Smarteye recommends that the system be given some time to "fill up" the buffer for this measure so that its reported values stabilize. Therefore,

removing data associated with the beginning of the runs had the added benefit of stabilizing the quality measure. Unless specified otherwise, analyses were conducted on only POG data that were associated with a gaze direction quality of 0.7 or greater. Regretably, data loss and insufficient data quality cannot be considered random errors. Situations in which pilots gaze was extreme (downward or to the side) were more likely to result in lost or poor quality data. As such, data from the EFB_Aft condition was disproportionally affected.

4   Analysis

It is not unusual, in experiments with relatively few participants and variables in which a lower bound is capped (e.g., response times), to obtain non-normal data, and for human individual differences in complex tasks to produce data which fails equal variability assumptions. To permit analysis of the mixed, repeated measures model but mitigate the effects of these departures from standard Analyses of Variance (ANOVA) assumptions, longitudinal data analysis was employed in which the covariance structures that account for the presumed intercorrelations attendant in repeated measures designs were specified. Compound symmetry (assuming heterogeneous variances and constant correlations among repeated measures) and diagonal (assuming heterogeneous variances and uncorrelated repeated measures) covariance structures were applied, and the model with the lowest Akaike criterion (i.e., is best fitting with the fewest parameters) selected. Robust (Huber-White) corrections for standard errors were employed to counteract distortions introduced by data with unequal variances. Satterthwaite adjustments were applied to the degrees of freedom to accommodate for unequal variances and unbalanced (missing) data in conditions. This type of adjustment is also are helpful in experiments with fewer data points. Finally, to counteract the effect of inflated alpha with multiple post-hoc comparisons, p-values were Sidak-adjusted. Type III sums of squares were used to ensure that order of terms in the model do not affect significance levels, and to protect against imbalance that may occur due to missing data.

Where data are interval count data, these were analyzed according to a Poisson distribution (rather than Normal) with a logarithmic link function. Particulars of other required transformations are discussed within the section for each data type. The factors considered in models differ by data type and therefore, are described in the following sections.

This study is one in a line of investigations intended to converge upon appropriate flightdeck interface designs to support IM. As such, a liberal alpha (p=0.10) was used for significance testing. This increases the power of tests with the intent to discover potentially significant findings with this relatively small sample size study, acknowledging the known risk that type 1 errors (assigning significance to effects that may not be so) may exist. The discussion of significant results is accompanied by p-values so the reader with a more conservative approach can discount less significant results. Analyses were conducted with SPSS statistical software package, v21.1.

4.1   Simulation Data

Simulation data (aircraft performance and position, control inputs, IM events) were collected at 5Hz during each run. Timestamped parameters were used to assess vertical excursions from the VNAV path (outside of 400 feet vertically), and speed excursions from the intended IM profile (over seven knots). The maximum excursion observed from the time a commanded speed occurred until 19 seconds later was used assess the impact of commanded speed changes on path and speed profile adherence. Inspection of raw data indicated that data for vertical excursions from the path were skewed right and so were lognormal transformed for analysis. These data were analyzed for effects of Aviation Condition, Notification Method, Pilot Flying (which side of the flightdeck), and all two-way interactions of these factors. The covariance matrix was defined by repeated measures on Avionics Condition and Pilot Flying.

Simulation data also contained data that permitted calculation of count data, including: the number of out-of-speed-conformance (speed deviations) that occurred, as well as the number of IM reminders that occurred. These data were analyzed according to a Poisson distribution with log link function. These data were analyzed for effects of Aviation Condition, Notification Method, Pilot Flying, and all two-way interactions of these factors. The covariance matrix was defined by repeated measures on Avionics Condition and Pilot Flying.

Crew response time was defined as the time from a commanded speed to the first instance of a change in the speed entered into the MCP. These data were analyzed according to a model similar to the above, but also included a repeated measure term to account for multiple observations per run (i.e., crew responses to each commanded speed).

4.2   Post-Run Questionnaire Data

Post-run questionnaire items were rated on a Likert 9-point scale in which low ratings correspond to "Completely Disagree" and high ratings correspond to "Completely Agree." Post-run questionnaire data were analyzed for effects of Aviation Condition, Notification Method, Pilot Flying, and all two-way interactions of these factors. The covariance matrix was defined by repeated measures on Avionics Condition and Pilot Role (whether the pilot had participated as PF or PM during the run). Frequency tables for each ordinal dependent variable provide the basis for characterizations of scale use. Descriptive statistics and figures include both mean

and median estimates of central tendency because while parametric statistics were employed, ratings scales can be considered ordinal data. The reader can see that these are similar in most cases, providing confidence for the use of the parametric approach. In addition to assessing Avionics Configurations using this instrument, item reliability analyses were conducted to determine the best subset of questions that would be useful in subsequent FIM studies.

4.3   Post-Experiment Questionnaire Data

Post-experiment questionnaire data were typically analyzed for effects of Notification Method, as all crews experienced the Avionics Condition. Because both pilots fly half of the scenarios as PF and half as PM, these distinctions were not appropriate for assessment of these summary data. In some cases, the same question was asked for the different Avionics Conditions, and this term was included in analysis models. Post-experiment questionnaire data was analyzed similarly with mixed generalized linear models as above.

4.4   Oculometer Data

Statistics were calculated based on Gamma distributions using a log link function, as most data were defined by non-negative values, and all distributions were positively skewed. Models included main effect terms (Avionics Condition and Notification Method), and the two-way interactions of these main effects. For some measures, each pilot provided data (e.g., Noticing Time); whereas for others, the crew served as the experimental unit (e.g., Minimum Noticing Time). When the experimental unit was a pilot, the Role (Pilot Flying (PF) or Pilot Monitoring (PM)) and interactions of Role with Avionics Condition and with Notification Method were included in analyses.

5 Results

In expressing these results, the term *Avionics Configuration* refers to the interaction of an Avionics Condition and a Notification Method as experienced by crews. All results are assessed for significance at alpha=0.10, but p-values are provided for the reader who choses to consider more stringent criteria.

5.1 Simulation Data

Simulation data were taken to describe vertical excursions, speed excursions and the number of out-of-conformance incidents, the number of reminders that occurred, and response times to dial in commanded speeds. Lateral excursions from path were not considered due to the influence of turns in scenarios.

Vertical Excursions

Crews were asked to maintain vertical path accuracy within 400 feet of the path. Vertical deviation error was calculated as excursion distance beyond this point, on either side, during a 19-second period following speed commands. Over all runs, the average maximum deviation was 272.41 feet out of tolerance (outside 400'). Approximately 31% of all runs were within tolerance, and 90% of the runs had max deviation of less than 890 feet out of tolerance, the largest being 1648 feet out of tolerance. Figure 11 shows the mean maximum vertical deviation over 400 feet following command speeds. The interaction between Notification Method and Avionic Condition (p<0.025) was significant. The only pairwise comparison that reached significance for this interaction showed that the VVV method was associated with more vertical excursions than the AAA method for the EFB_Aft condition. There was a significant interaction between Notification Method and Pilot Flying (p=0.072); whereby, when experiencing the AAA

method, pilots flying from the left seat had more exaggerated vertical excursions than those flying from the right seat.
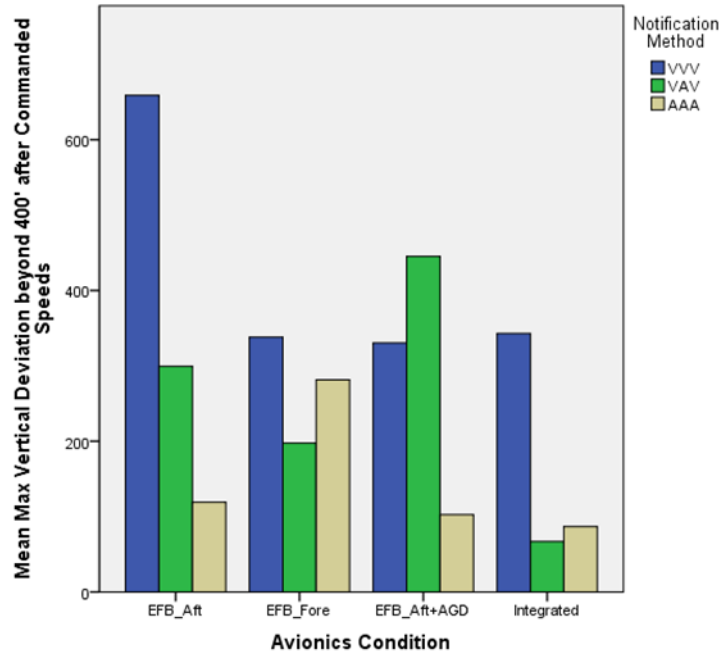
Figure 11. Averaged Maximum Vertical Excursion beyond 400', after Commanded Speeds.

Speed Excursions & Out-of-Conformance Incidents

Max speed excursions from the IM profile speed were calculated for each run by taking the maximum value of speed error in the 19 second windows following each commanded speed change (Figure 12). Over all runs, crews averaged 14.6 knots off profile (or 7.6 knots out of conformance as they were asked to maintain indicated air speed within 7 knots of the IM speed profile), and the standard deviation of these data was 6.22 knots. Avionics Condition affected the degree to which speed profile adherence was problematic (p=0.016). The EFB_Aft condition showed greater speed excursions than those in the EFB_Aft+AGD (p=0.069) and the Integrated (p=0.024) conditions. The interaction of Avionics Condition and Notification Method was also significant (p=0.020), and the significant pairwise comparisons show that this main effect is primarily due to much larger speed excursions when using the EFB_Aft condition with the VVV method than either VAV (p=0.016) or AAA (p=0.007).

The number of out-of-conformance indications was analyzed as another measure of how well crews were able to control aircraft speed to the intended speed profile. Figures 13 and 14 show the mean and median number of times pilots received speed profile conformance deviation indicators. The significant interaction term (p=0.011) reveals significant differences in Notification Method for both the EFB_Aft and Integrated conditions that follow the same pattern; crews with the VVV Notification Method had two, to three, times more instances of being out of conformance with the speed profile than did crews with either the VAV condition (for EFB_Aft, 0.089; for Integrated, p=0.058) or the AAA condition (for EFB_Aft, 0.089; for Integrated, p=0.023). Main effects are consistent. Over all levels of Avionics Condition, Notification Method was a significant factor (p=0.009), with the primary difference being that the VVV method was associated with more out-of-conformance indications than the AAA method (p=0.049). Over all levels of Notification Method, Avionics Condition was significant (p<0.001); where crews in the Integrated condition showed significantly more out-of-conformance indications than any of the other conditions (EFB_Aft, p=0.001; EFB_Fore, p=0.002; EFB_Aft+AGD, p=0.005).
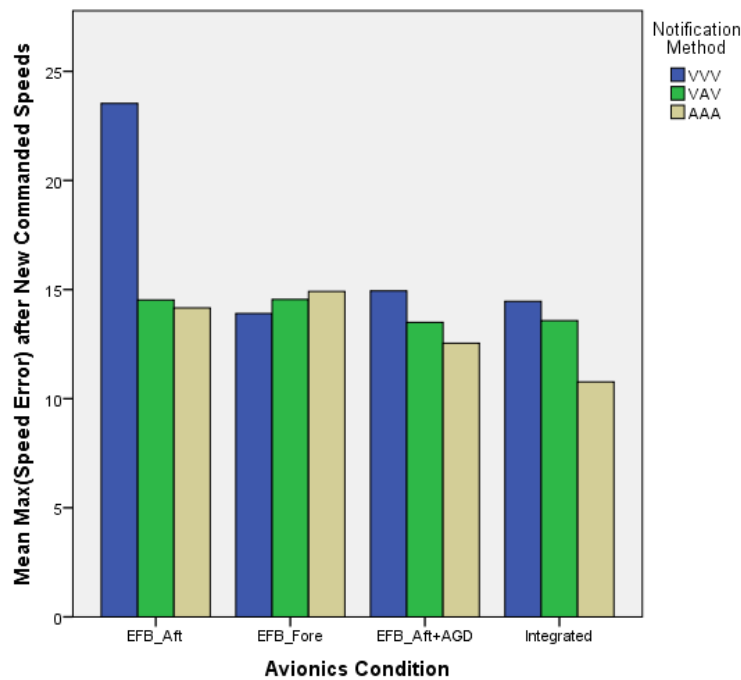

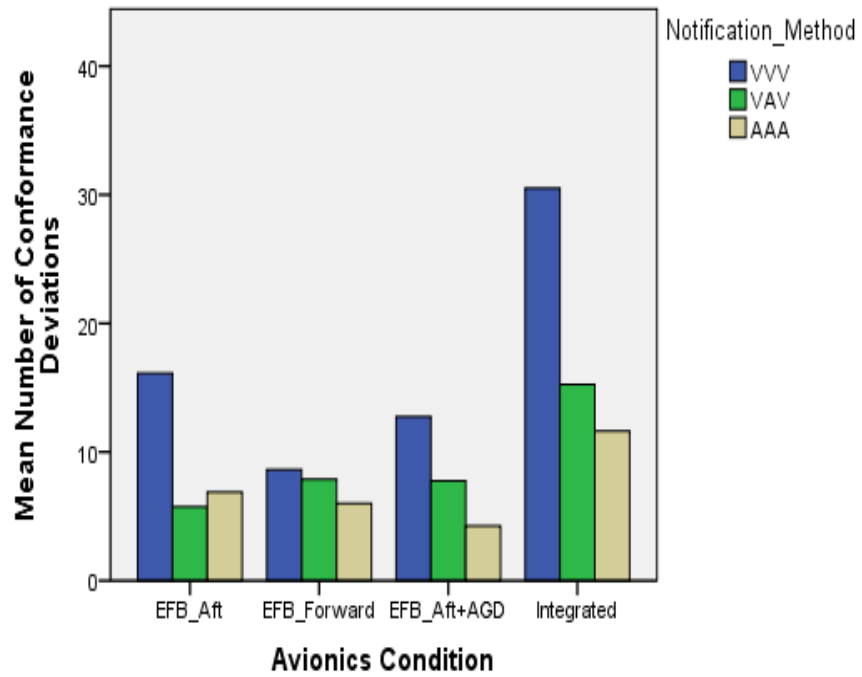**Figure 12. Averaged Maximum Speed Excursion after Commanded Speeds.**

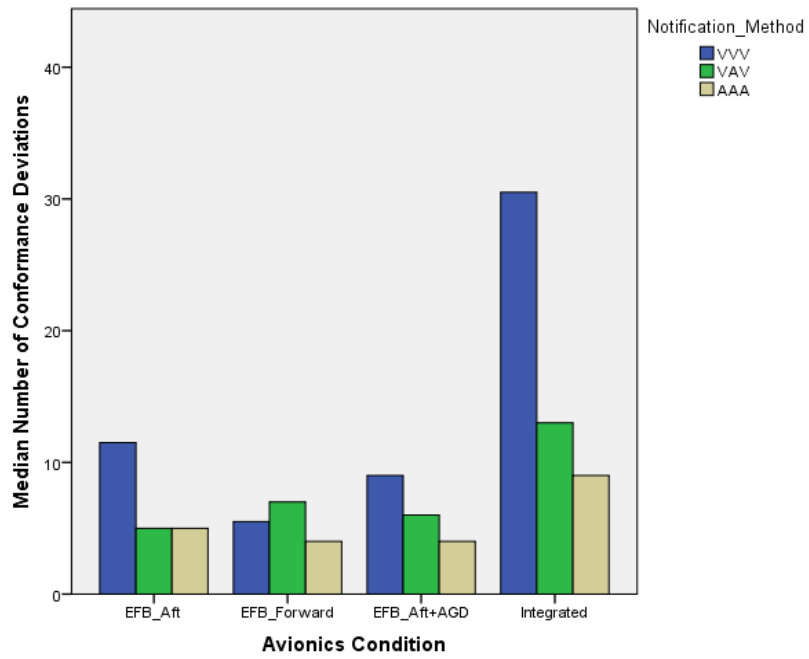**Figure 13. Mean Number of Conformance Deviations.**

Figure 14. Median Number of Conformance Deviations.

Attentiveness to IM Events: Reminders and Response Times

In addition to the out-of-conformance events, data were analyzed to ascertain crew's responsiveness to target speeds, and to determine how many of these were missed, i.e., how many times a reminder event occurred.

The maximum number of reminders over all runs, was seven; with a mean of 1.75, and a median of one. Figures 15 and 16 show the mean and median number of reminders experienced per experimental conditions. Both Avionics Condition (p=0.002) and Notification Method (p<0.001) were significant for missing initial speed targets and triggering these reminders. Inspection of pairwise comparisons show that when crews used the EFB_Fore condition they received significantly fewer reminders (and so entered speeds on the first indication) than with either the EFB_Aft (p=0.001) condition or the EFB_Aft+AGD (p=0.094) condition. Crews that used the AAA Notification Method had significantly fewer reminders than either those using the VVV (p=0.003) or the VAV method (p=0.075). The VAV method had significantly fewer reminders than the VVV method (p=0.075). The average number of reminder indications for VVV was about three times that of the VAV method. The average for the VAV method was about three times that of the AAA method.

To characterize responsiveness to commanded speeds, a response time was calculated from the time the commanded speed was issued to the time for one of the crew members to change the entry in the MCP speed window. Figure 17 shows the means for these data. Median response times are between 2.6 and 5.2 seconds, but for some conditions maximum values are closer to a half-minute. As reaction time data have a lower bound, but no upper bound, data were log transformed for analysis.

Transformed response time was significantly affected by Avionics Condition (p<0.001), Notification Method (p=0.002), and the interaction between Pilot Flying and Notification Method (p<0.001). For the Integrated and the EFB_Fore conditions, crew response times were significantly shorter than for either the EFB_Aft (for Integrated, p<0.001; for EFB_Fore, p=0.002) or EFB_Aft+AGD (for Integrated, p=0.003; for EFB_Fore, p=0.003). For the VVV Notification Method, crew response times were longer than either the VAV (p=0.033) or AAA (p=0.003) methods. When using the AAA method, response time was greater when the pilot flying was in the right seat (p<0.001), and when using the VAV method, response time was greater when the pilot flying was in the left seat (p=0.065).
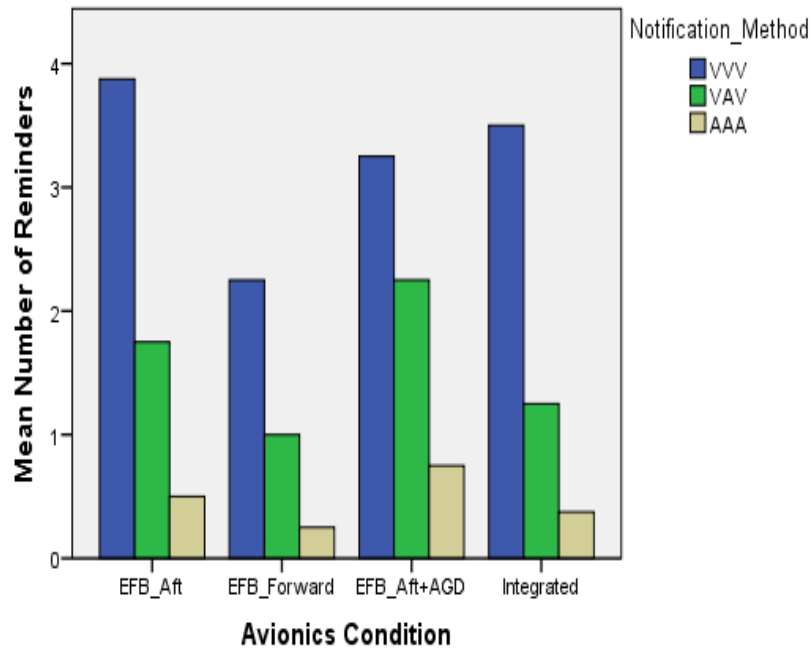
Figure 15. Mean Number of Reminders.
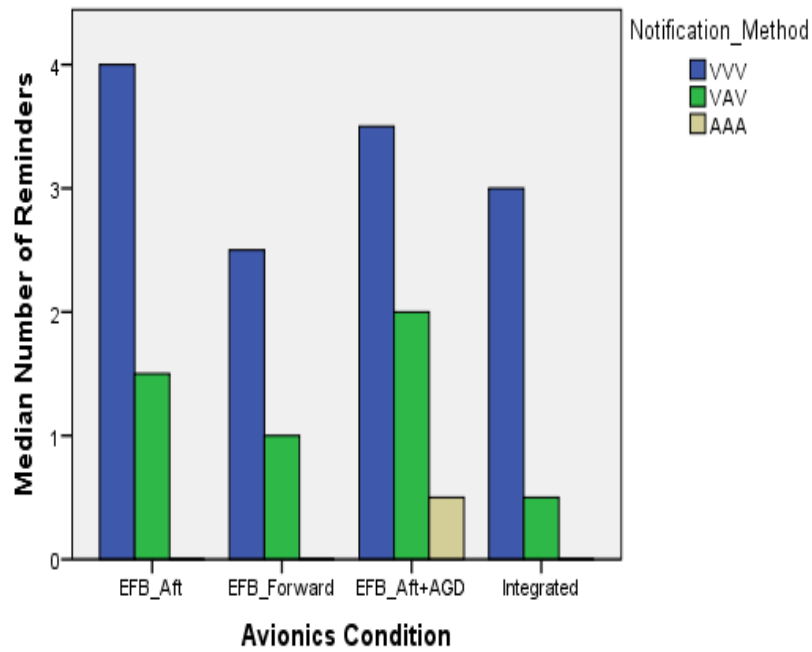
Figure 16. Median Number of Reminders.
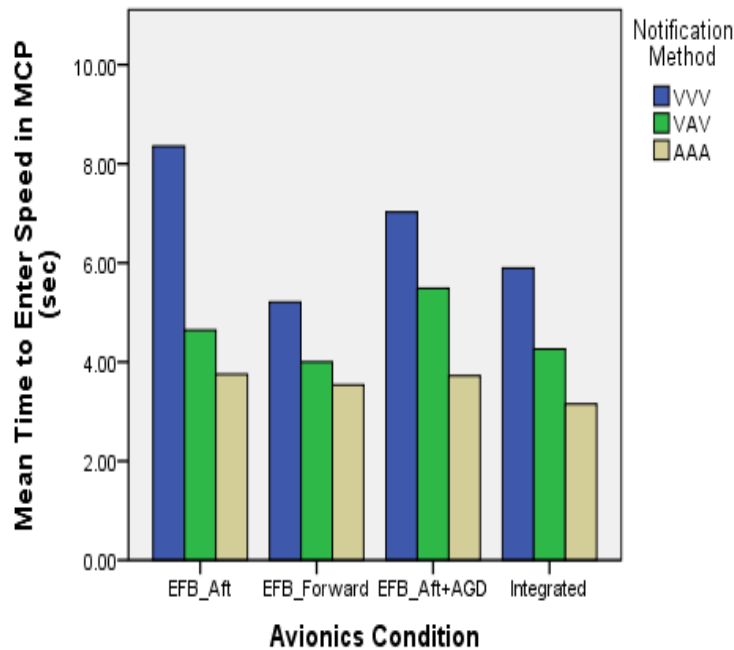
Figure 17. Mean Time to Enter Speed in MCP.

5.2 Post-run Questionnaire Data

The post-run questionnaire data are presented in the following sections. References to Likert items are denoted "Q##," where the item number is provided (and as is denoted on the questionnaire in Appendix D). In a few cases, pilots failed to provide a rating on a post-run questionnaire item. Typically only one response was missing (1% of the data) and at most, 3 ratings (1.6%) were missing. The maximum missing data was from Q16 and the composite score associated with Speed Deviation Awareness. Models for these data included terms for the Avionics Condition, Notification Method, the Role of the respondent during the particular run (PF or PM), and all two-way interactions. Repeated measures were specified, and the covariance structure applied to, the interaction of Role and Avionics Condition.

Workload

The post-run questionnaire included the Modified Cooper-Harper (MCH) scale (Cooper & Harper, 1969) (where pilots were asked to self-assess their average and and their peak workload); the NASA Task Load Index (TLX) (Hart & Staveland, 1988) (where pilots rate

workload on mental demand, temporal demand, performance, effort, frustration subscales); and Likert scale items.

Modified Cooper-Harper Ratings

Pilots were asked to rate their average and peak workload after each run using the MCH scale, from 1 (easy) to 10 (impossible to complete). Pilots' mean and median ratings of MCH workload ratings are provided in Figures 18 (for average workload in a run), and 19 (for peak workload in a run). All subject ratings of average workload on this scale were less than or equal to 4, corresponding to the statement *"Moderately high operator mental effort is required to attain adequate system performance"*, and all ratings for peak workload were less than or equal to a five, corresponding to the statement "*High operator mental effort is required to attain adequate system performance".* However the vast majority of the data (99% for average workload, and 95.3% for peak workload) was rated three or less – meaning, at most, for these runs *"Acceptable operator mental effort is required to attain adequate system performance."* Ratings of average and peak workload using the MCH scale showed no significant effect of configuration factors (all p≥0.181), but showed a main effect of Role whereby PFs experienced both higher average workload (p=0.041) and peak workload (p=0.043).
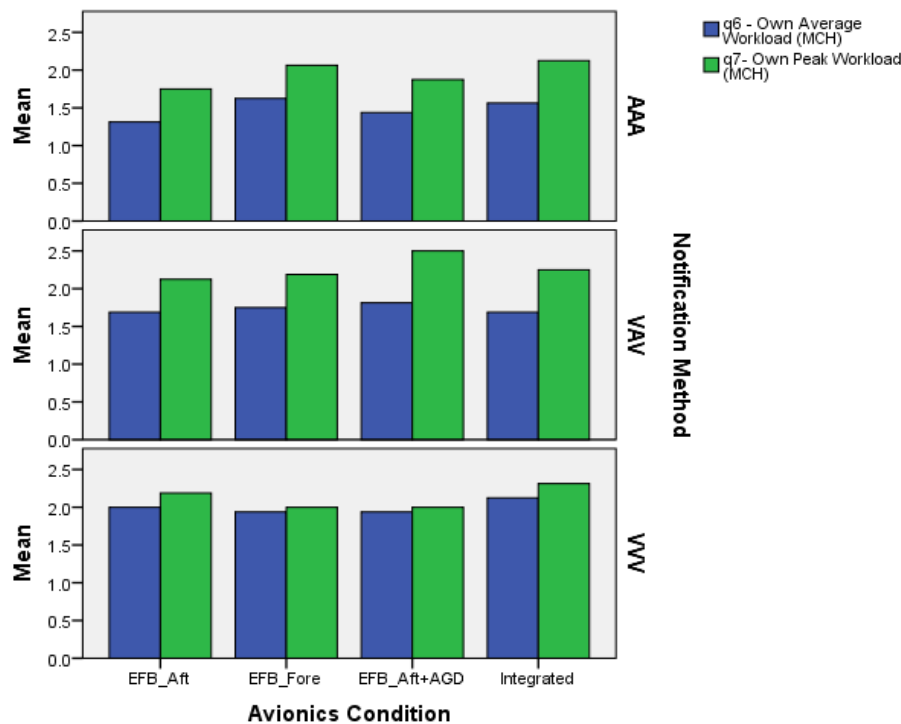
Figure 18. Mean MCH Average and Peak Workload Ratings (Scale ranges 1 to 10).
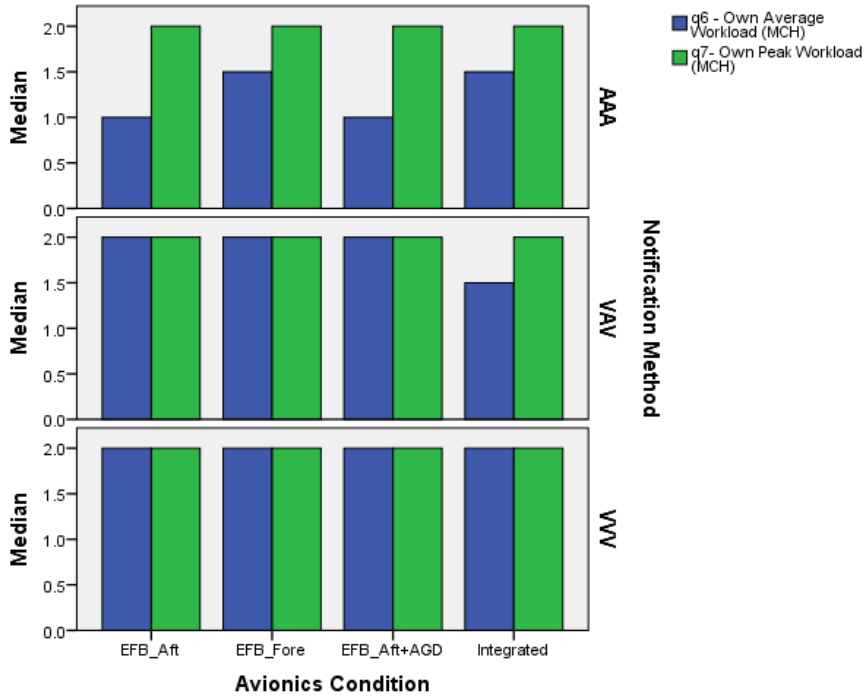
Figure 19. Median MCH Average and Peak Workload Ratings (Scale ranges 1 to 10).

NASA TLX Ratings

The NASA-TLX subscales (mental demand, temporal demand, effort, perceived performance, and frustration) provide ratings from 0 to 20, where high scores correspond to high perceived workload. A total TLX score was constructed by adding the scores on these five subscales, resulting in a range from 0 to 100.

Figures 20 and 21 show the mean and median ratings for NASA TLX scales. Analysis of frequency data showed low workload scores. For the mental demand subscale, just over 93% of the ratings were at or below the midpoint (10), and over 63% were five or less. For the temporal demand subscale, almost 95% of the ratings were at or below the midpoint (10), and just over 73% were six or less. For the perceived performance subscale, over 81% of the ratings were at or below a rating of nine (no midpoint ratings were provided), and just over 75% were five or less. Note that the perceived performance subscale is reverse-scored such that lower ratings equate to higher self-assessed performance. For the perceived effort subscale, almost 93% of the ratings were at or below the midpoint (10), and approximately 68% were five or less. For the frustrating subscale, just over 97% of the ratings were at or below the midpoint

40

(10), and just over 89% were five or less. For the totaled score (ranging zero to 100), almost 94% of the ratings were at or below the midpoint (50), and almost 69% were 25 or less.

Avionics Condition did not affect ratings on the TLX scales or composite scale (all $p \geq 0.257$). All the TLX subscales (and therefore the composite scale) showed a significant effect of Notification Method (mental demand $p=0.069$, temporal demand $p=0.052$, performance $p=0.072$, effort $p=0.029$, frustration $p=0.021$, total score $p=0.026$). For all these scales except the frustration subscale and the total scale, pairwise comparisons did not reach significance (all $p \geq 0.106$). Pilots' ratings indicated a significant increase in frustration with the VAV method over the AAA method ($p=0.027$). Also for the total TLX workload score, the VAV method had a significantly higher workload ($p=0.083$) than the AAA method. Pilots using the VVV method provided higher workload scores than those using the AAA method ($p=0.083$).

The temporal demand ($p=0.007$), perceived effort ($p<0.001$), frustration ($p=0.009$), and total TLX score ($p=0.006$) showed a significant main effect of Role; in all cases the PF rated workload higher than PM. In some cases, Role also interacted with configuration variables. Temporal demand ratings showed a significant interaction of Role with Notification Method ($p=0.036$); for the VAV method, ratings were higher for the PF than for the PM ($p=0.005$).

For the mental demand ($p=0.068$), performance ($p=0.006$), and effort ($p=0.035$) subscales, ratings were significantly affected by the interaction of Role and Avionics Condition, and this is also reflected in assessment of the total TLX score ($p=0.004$). For all these subscales, when using the EFB_Fore condition, pilots rated workload higher when they were PF {mental demand ($p=0.001$), performance ($p=0.076$), effort ($p=0.030$)}, and this interaction was reflected in the TLX total score as well ($p=0.003$).
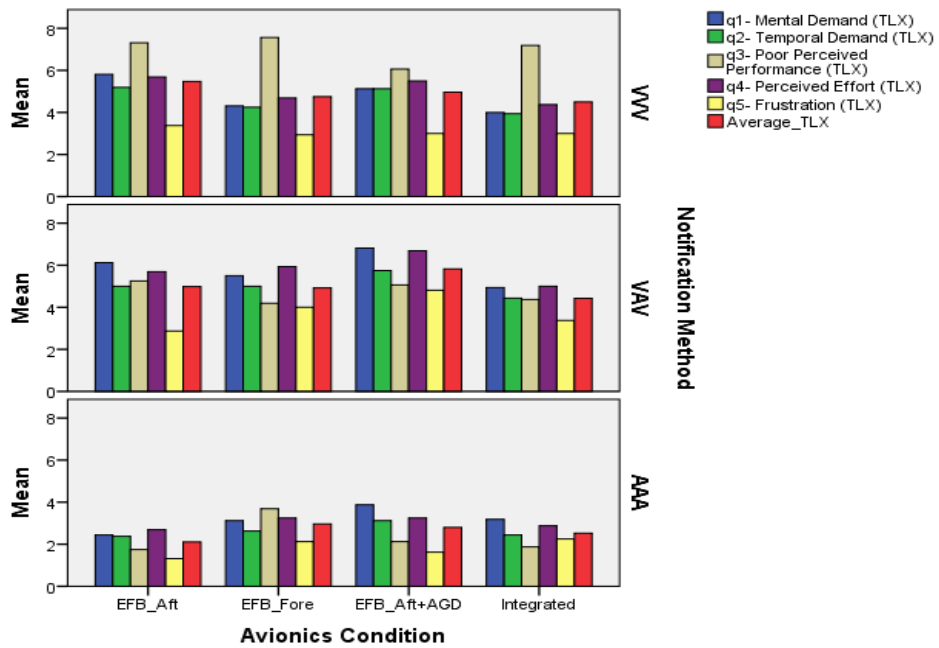
Figure 20. Mean NASA TLX Ratings (Scale ranges 0 to 20).
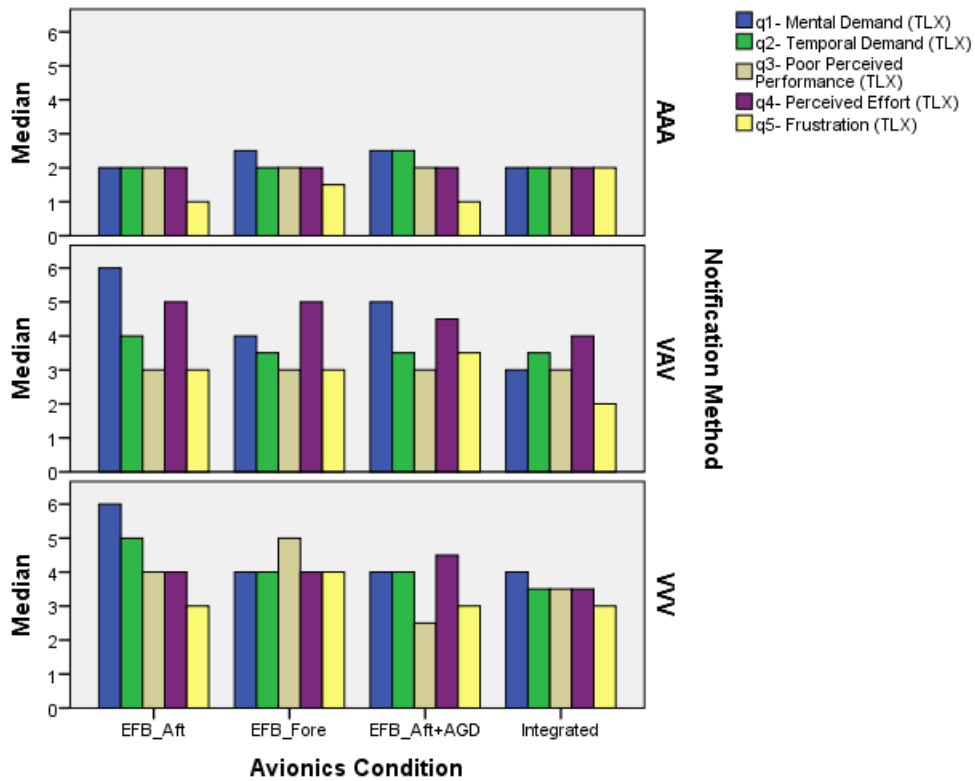
Figure 21. Median NASA TLX Ratings (Scale ranges 0 to 20).

Situation Awareness

Situation awareness (SA) was assessed using the three-point SA Rating Technique (SART) scale (Taylor, 1990), which has subscales pertaining to "*demand* on attentional resources," "*supply* of attentional resources," and "*understanding*." In addition, seven-point Likert items were used to differentiate among elements of situation awareness, in particular for: awareness of new speed targets (see speed awareness questions Q15, Q16, Q22, Q26 in Appendix D), awareness of adherence to the IM speed profile or deviations from IM speed (Q23, Q27), awareness of the TTF aircraft (Q21, Q35), and awareness of other aspects of flight (Q29). Composite scores were calculated for these constructs based on the average of responses for the constituent items.

SART Scores

Each subscale ranges from 1 to 7. A rating of 7 indicates a less demanding situation for the supply and understanding subscales, and a more demanding situation when using the demand subscale. The SART score is defined by the equation {understanding – (demand – supply)} and results in a scale ranging from -5 to 13. SART scores were rescaled to range from 0 (low situation awareness) to 18 (high situation awareness).

For the demand scale, approximately 72% of the data was rated a 3 or lower. For the understanding scale, approximately 92% of the ratings were 5 or higher. For the supply scale, approximately 75% of the ratings were 5 or higher. Approximately 75% total scores were at or above a rating of 7. Figures 22 and 23 show the mean and median total SART scores.

Avionics Condition had a significant effect on overall SART score (p=0.067). Pilots rated situation awareness lower when using the EFB_Aft condition than when using the Integrated condition (p=0.090). Pilot Role significantly affected ratings for demand (p=0.005), understanding (p=0.043); and for both these, PF ratings were higher. Pilot Role also significantly interacted with Avionics Condition for overall SART ratings (p=0.031); revealing that for PF, situation awareness was lower with the EFB_Fore condition than the Integrated condition (p=0.70).

Four questions assessed participants' awareness of commanded speed changes: *"Avionics and annunciations provided in this scenario appropriately supported my ability to detect speed target onsets in a timely manner"* (Q26), *"I was aware of commanded speed changes within an appropriate time frame"* (Q15), *"I was able to implement the speed changes in the MCDU within an appropriate timeframe"* (Q16), and *"I was never surprised to notice that the speed target had changed and wondered how long ago it had"* (Q22). A composite score was constructed from the mean of these items. All items were rated as 9 point Likert items with a high score signifying strong agreement with the item's statement. Figures 24 and 25 show the means and medians associated with the individual questions and the composite score for awareness of commanded speed changes. Approximately 91% of the ratings for Q26 are positive (a rating of five or greater), 94% of the ratings for Q15, 97% for Q16, and approximately 94% of the ratings for Q22 are positive.

All speed awareness items, and the related composite score showed significant main effects for both Avionics Condition (all p≤0.026) and Notification Method (all p≤0.030). In most cases, these ratings showed significantly lower scores for the EFB_Aft condition than the EFB_Fore condition (all but Q16, p≤0.092); and in all cases, significantly higher scores for the AAA method than the VVV method (all p≤0.008). Q26 ratings were also significantly higher for both the EFB_Aft+AGD condition (p=0.041) and the Integrated condition (p=0.033) than the EFB_Aft condition. In fact, for Q26, all other conditions were rated more highly than the EFB_Aft condition (all p≤0.041), and the Integrated condition was rated more highly than the EFB_Aft+AGD condition (p=0.098). Q15 and Q22 both also revealed significantly higher ratings for the EFB_Fore condition than the EFB_Aft+AGD condition, (p=0.021 and p=0.063, respectively). The composite score revealed the same effect with respect to the EFB_Aft and EFB_Fore conditions, and also higher ratings for the Integrated condition than the EFB_Aft+AGD condition (p=0.094).

Question 15 and composite score ratings also reflected a significant interaction between the Avionics Condition and the Notification Method (p=0.011 and p=0.065, respectively). For the EFB_Aft, EFB_Fore and Integrated conditions, ratings on both scales reflect the Notification Method main effect; whereby the AAA method was rated higher than the VVV method (all p≤0.044). However Q15 also had significantly higher ratings for the VAV method than the VVV method for the Integrated condition (p=0.066), and composite ratings show that for the EFB_Aft condition, ratings with the AAA method are significantly higher than with the VAV method

(p=0.044). In addition to showing the same main effect for Notification Method (AAA higher than VVV) as other constituent questions and the composite (all p≤0.008), ratings on Q22 were higher for pilots using the AAA method than those using the VAV method (p=0.072).

No significant effects of Role (all p≥0.252) or interactions of this term with Avionics Condition or Notification Method were observed (all p≥0.102).
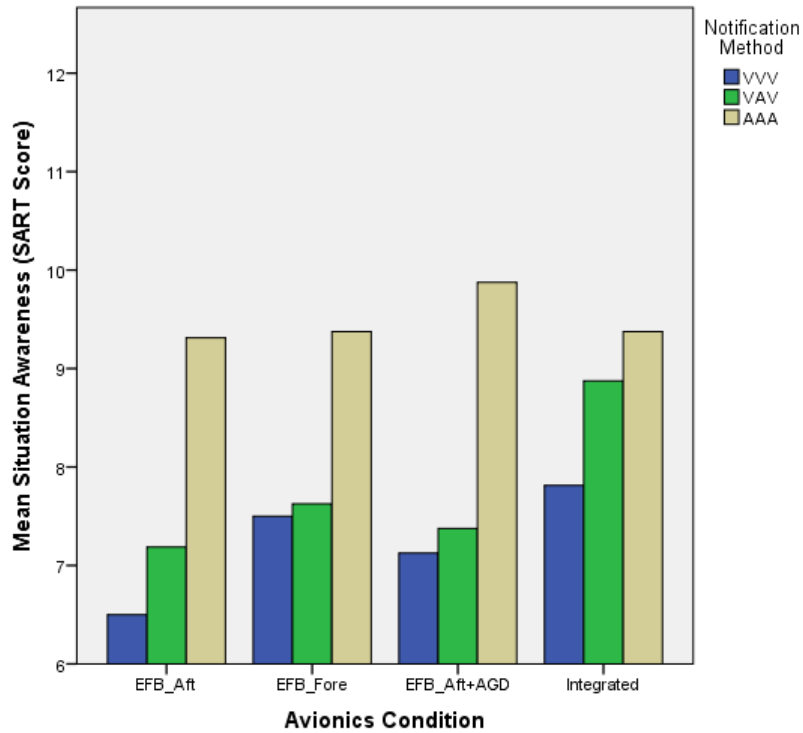
Figure 22. Mean SART Scores (Scale ranges 0 to 18).
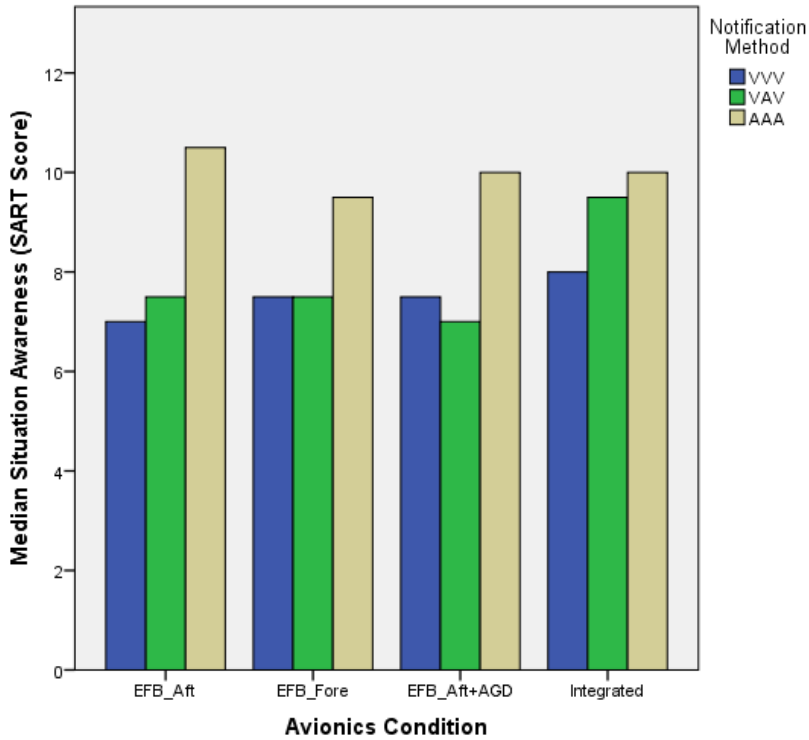
Figure 23. Median SART Scores (Scale ranges 1 to 18).

Figure 24. Mean Speed Onset Awareness Ratings (Scale ranges 1 to 7).
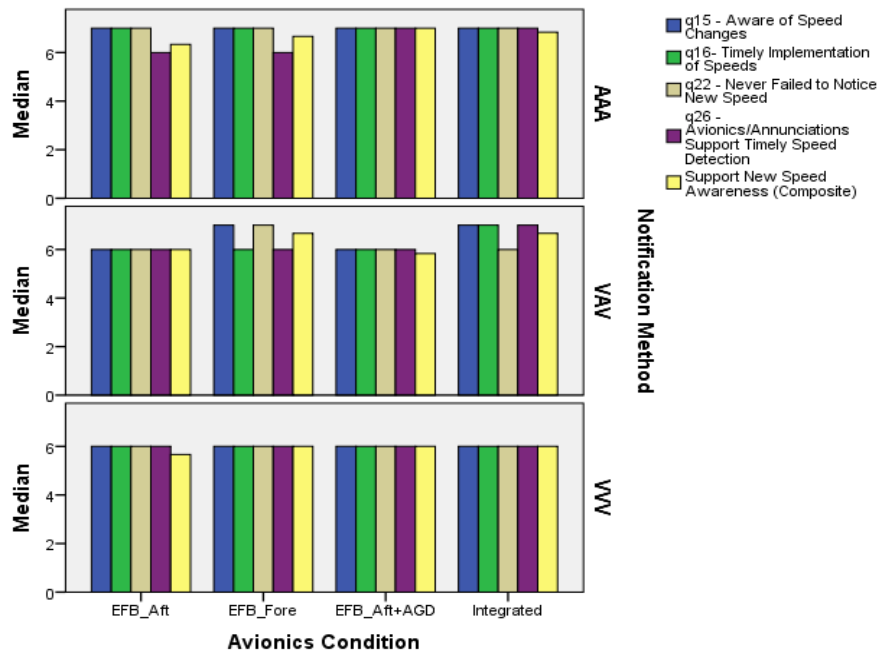
Figure 25. Median Speed Onset Awareness Ratings (Scale ranges 1 to 7).

Speed Profile Deviation Awareness

Two questions were used to assess pilot's awareness of their speed relative to the IM speed profile: *"The avionics and annunciations provided in this scenario provided appropriate support for detecting deviations from speed profile in a timely manner"* (Q27), *and "I was aware of speed profile conformance throughout the scenario."* (Q23). A composite rating was constructed from the averages of these items. Figures 26 and 27 show the means and medians associated with these ratings. Analysis of these data indicates that subjects generally felt comfortable with speed profile awareness. Approximately 81% of ratings on Q27 were positive (five or greater). Approximately 91% of ratings on Q23 were positive. Finally, over 91% of the composite ratings were positive. The minimum score on the composite question and Q23 was four. Q27 had 9.4% of ratings, four or less.

All questions and the composite question reflected a significant effect of Notification Method on awareness of speed profile (all p≤0.036). All these scale items showed one common effect: ratings for the AAA method were significantly more positive than those in the VVV method (all p≤0.029). Q27 ratings were significantly higher for the AAA method than for the VAV method (p=0.033), and this finding emerged in the composite score results as well (p=0.092).

Q27 was the only scale item that was sensitive to effects of Avionics Condition (p=0.028). Ratings on this question showed an advantage of all other conditions over the EFB_Aft condition (vs EFB_Fore, p=0.078; vs EFB_Aft+AGD, p=0.044; vs Integrated, p=0.047).
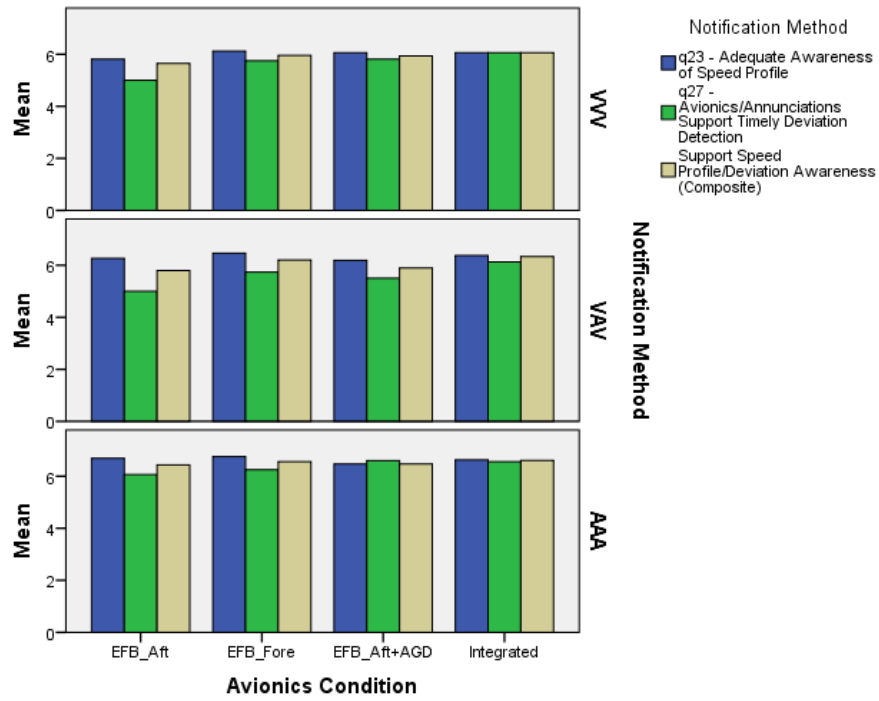
Figure 26. Mean Speed Profile Awareness Ratings (Scale ranges 1 to 7).
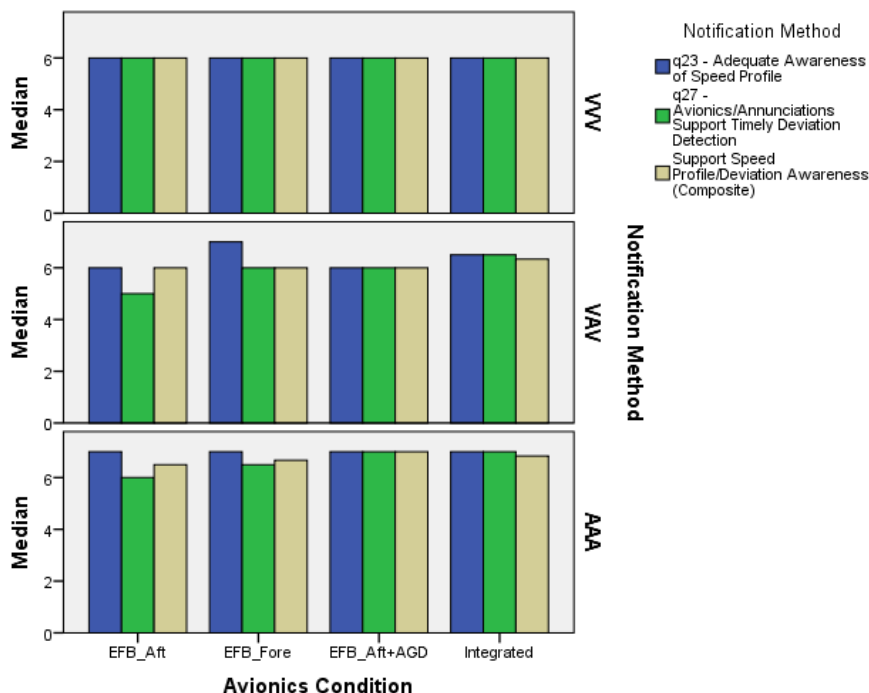
Figure 27. Median Speed Profile Awareness Ratings (Scale ranges 1 to 7).

TTF Awareness

Two questions addressed pilots' awareness of the TTF: *"I maintained adequate awareness of my lead aircraft throughout the scenario"* (Q21) and *"I was comfortable with the location of my target aircraft throughout the run"* (Q35). A composite score was calculated based on the average of these items. Generally, ratings were positive on these scale items. Approximately 93% of the ratings for Q35 and 89% of the ratings for Q21 were positive (a score of five or greater). Figures 28 and 29 show the means and medians of these ratings. Ratings associated with TTF awareness (Q21) showed no significant main effects for the Avionics Condition factor (all p≥0.512). Q35 (p=0.070) ratings showed a significant main effect for Notification Method, whereby ratings for the AAA method are higher than those for the VVV method (for Q35, p=0.089). While the composite score also indicated this main effect (p=0.079), no pairwise comparisons were significant (all p≥0.100).
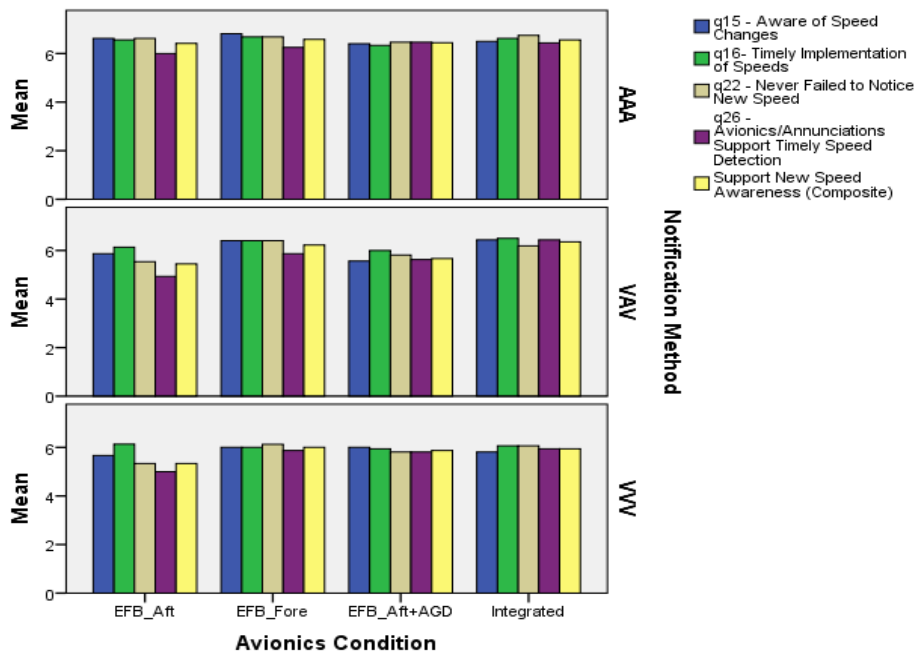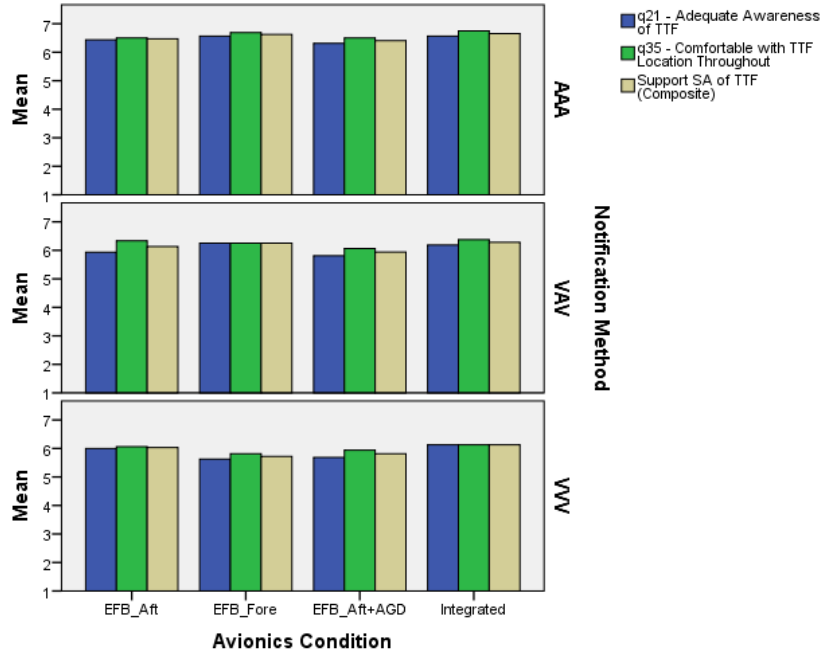
Figure 28. Mean TTF Awareness Ratings (Scale ranges 1 to 7).
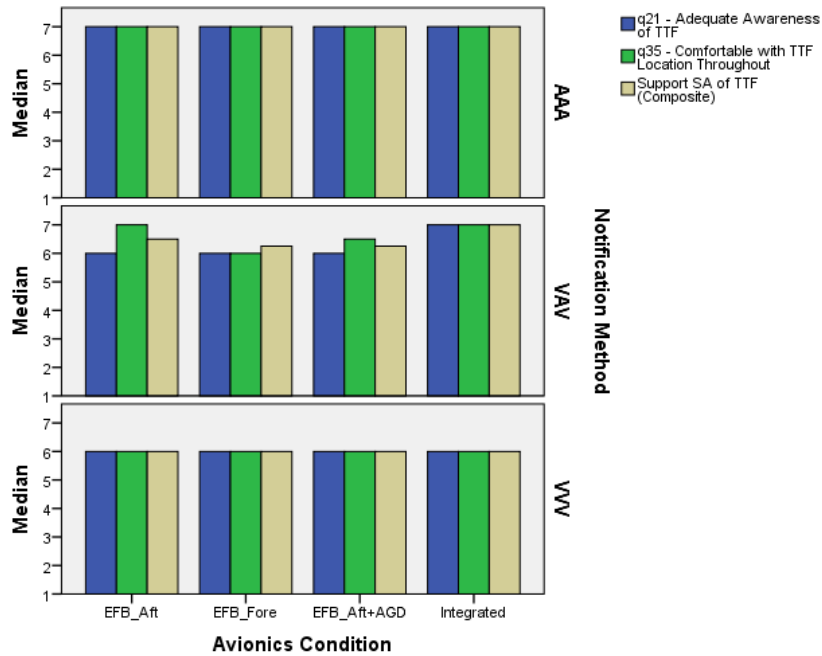
Figure 29. Median TTF Awareness Ratings (Scale ranges 1 to 7).

General Awareness

While the focus of this study is to assess display configurations' ability to support IM, it is important to know the impact of these operations on SA for other aspects of flight. Ratings for Q29 *"Time required for IM tasks did not detract from having appropriate SA for other aspects of flight"* addressed this concern. Generally, these ratings indicated low impact on building general SA; approximately 96% of the data were positive (five or greater). Figures 30 and 31 show the mean and median rating for this question, respectively. Ratings on this question were significantly impacted by Notification Method ($p=0.060$), Role ($p=0.006$), and the interactions of the Avionics Condition with both Notification Method ($p=0.028$) and Role ($p=0.059$). The only significant avionics configuration pairwise interactions show ratings for the AAA method to be significantly higher than the VAV method ($p=0.026$) and the VVV method ($p=0.002$), but only when using the EFB_Aft condition. The only aspect that is revealed in the main effect for Notification Method is the paired comparison between the AAA method and the VVV method ($p=0.051$), which is likely derived from the aforementioned interaction. A main effect of Role is significant ($p=0.006$), where by ratings for PM were higher than for PF; but this trend is only revealed in pairwise comparisons for the EFB_Aft ($p=0.085$) and Integrated ($p=0.024$) conditions.
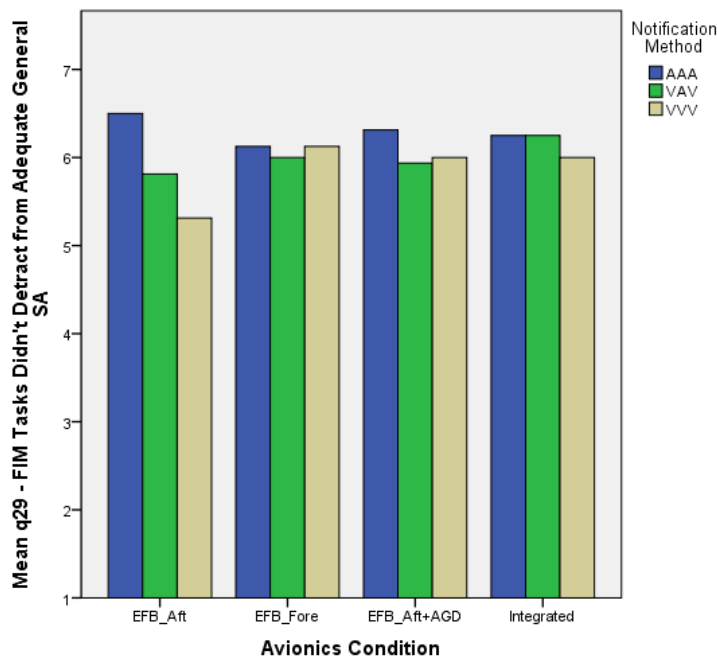
Figure 30. Mean General Awareness Ratings (Scale ranges 1 to 7).

51
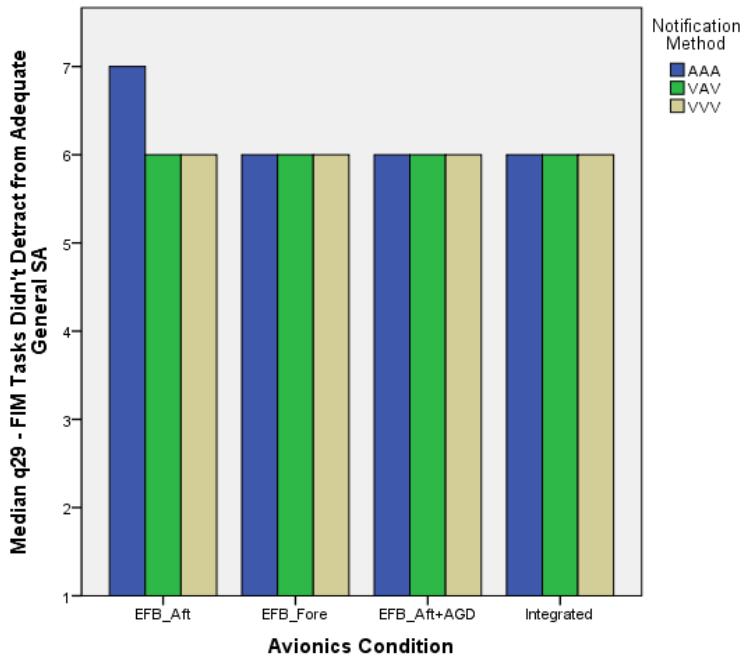
Figure 31. Median General Awareness Ratings (Scale ranges 1 to 7).

Usability

Three questionnaire items addressed general usability issues: *"The avionics and annunciations provided in this scenario were not overly distracting"* (Q24), *"Use of the avionics and annunciations provided in this scenario are acceptable for conducting IM operations"* (Q25), and *"I did not find the spacing tool frustrating during this scenario"* (Q34). A composite score was developed based on ratings from these three items. Ratings were generally on the positive (larger) ends of the 9 points scales: approximately 90% of the ratings were positive for Q24, approximately 90% were positive for Q25, and approximately 93% were positive for Q34. On the composite scale, approximately 90% were positive. Figures 32 and 33 show the mean and median rankings for these items.

Avionics Condition significantly affected ratings pertaining to distraction (Q24) and operational acceptability (Q25) ($p=0.003$, $p=0.001$, respectively), and this effect emerged on the composite score as well ($p=0.008$). A weakly significant interaction of configuration terms ($p=0.092$) showed that when using the EFB_Aft condition, the AAA method was less distracting than the VAV method ($p=0.015$); and for the Integrated condition, the AAA method's had significantly higher ratings (less likely to be distracting) than the VVV method ($p=0.098$). A main effect of

Avionics Condition for this item indicated that the Integrated condition was less distracting than either the EFB_Aft (p=0.016) or the EFB_Fore (p=0.022) conditions, and that the EFB_Aft+AGD was less distracting than the EFB_Aft condition (p=0.016). For Q25, general acceptability, all conditions were rated higher than the EFB_Aft condition (all p≤0.064), and the Integrated condition was rated higher than all other conditions (all p≤0.074). In addition, Notification Method was significant (p=0.047), and there was a weakly significant (p=0.098) interaction between this and Avionics Condition. The main effect shows ratings of acceptability higher for the AAA method than the VVV method, and this is found in particular in the interactions for the EFB_Aft+AGD condition (p=0.013). For the EFB_Aft without the AGD, the AAA ratings proved higher than the VAV ratings (p=0.016). Question 34 showed a significant effect of Notification Method (p=0.037), in which a pairwise comparison between ratings for AAA and VVV methods was significant (p=0.019) and means indicate the AAA method less frustrating. While these ratings show significant differences, it is important to note that the median differences show only one or two scale point differences, and that ratings were relatively high for all configurations.

The Role factor was not significant (all p≥0.297), and neither were any of the interactions of this factor with avionics configuration factors.
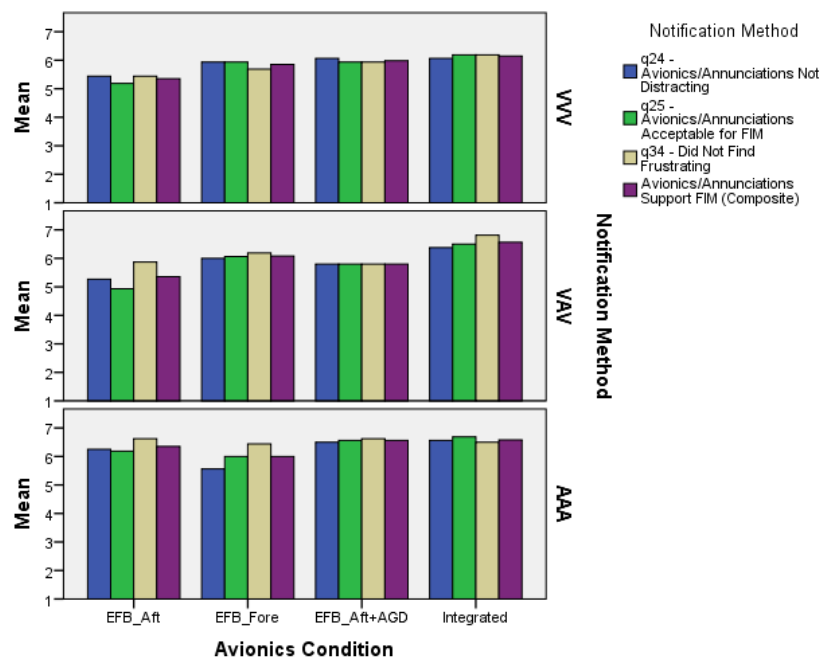
Figure 32. Mean Usability Ratings (Scale ranges 1 to 7).
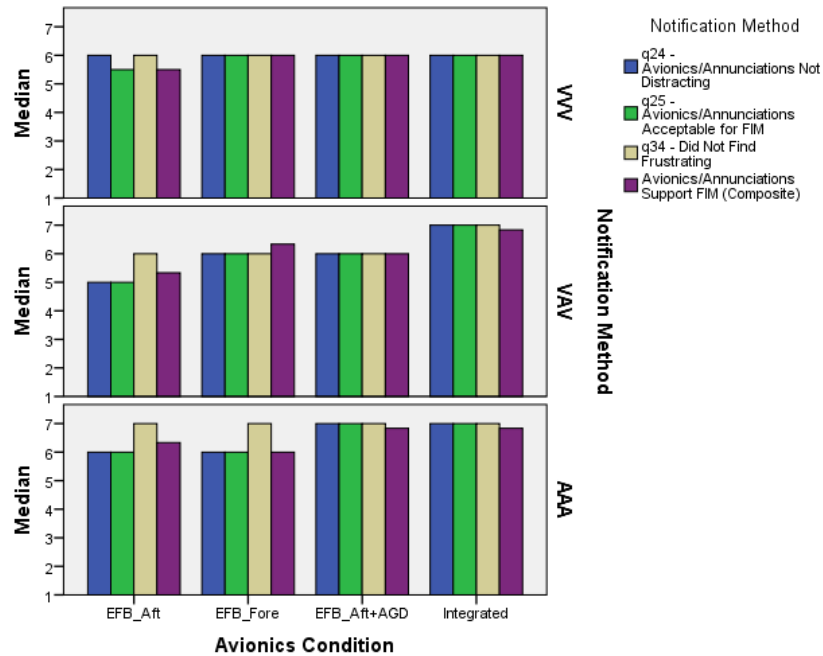
Figure 33. Median Usability Ratings (Scale ranges 1 to 7).

Scenarios / Operational Acceptability

The scenarios were designed to be realistic for KDFW arrivals. The post-run questionnaire contained several questions that addressed the degree of realism of the scenarios. Table 1 shows the questions, percent data scoring a five or greater, median and mean. Figures 34 and 35 show the mean and median ratings for these questions.

In the majority of cases, ratings were positive and consistent with the expectation that these factors would be invariant to avionics configuration manipulations. Questions 18, 19, 30, 33, and 36 showed significant effects of experimental manipulations associated with some aspect of Avionics Configuration. Avionics Condition was a significant main effect for Q18, Q19 and Q30 (all $p<0.04$). No significant pairwise comparisons were found for Q19 (whether speeds required an uncomfortable level of speedbrakes. For Q18, pairwise comparisons among Avionics Conditions, show that ratings after experiencing a run with the EFB_Aft condition were marginally better then when experiencing it with the EFB_Aft+AGD condition ($p=0.064$). For Q30, Avionics Condition significantly interacted with Notification Method; pilots who experienced the EFB_Fore condition with the VAV method provided, on average, higher ratings than those who had the VVV method ($p=0.003$). For Q33 ("… comfortable flying the commanded

speeds...") and Q36 ("Apart from IM operations... (workload was typical)..."), Notification Method was significant (all $p≤0.053$), and pairwise comparisons revealed differences, and means show direction, indicating pilots with the VAV method were more comfortable, than those with the VVV method (all $p≤0.094$). Some of these questions were also affected by the Pilot Role factor (Q17, $p=0.089$; Q18, $p=0.031$; Q20, $p=0.001$). In all cases, ratings from the pilot in the PM role were higher.

Table 1. Post-Questionnaire Items for Scenario / Operational Acceptability

Q#	Questionnaire Item (Strongly Disagree - Strongly Agree)	%Data≥ 5	Mean	Median
17	"The commanded speeds were operationally acceptable and appropriate"	95.9	6.30	6.00
18	"The frequency of IM speed commands was acceptable throughout the scenario"	96.9	6.32	6.00
19	"The speeds given did not require an uncomfortable level of speedbrake use"	87	5.85	6.00
20	"IM speeds and their timing are consistent with expectations for this type of arrival"	93.7	6.18	6.00
28	"The events I experienced in this scenario are operationally realistic"	97.4	6.37	6.00
30	"In this scenario, I thought commanded speeds were safe and appropriate to fly"	99	6.47	7.00
31	"I received IM speeds when in the process of completing other critical tasks"	72.3	5.05	6.00
32	"At no point in this scenario did I feel that that the commanded speed conflicted with other available information"	95.3	6.18	6.00
33	"At all times in this scenario, I felt comfortable flying the commanded speeds"	97.4	6.43	7.00
36	"Apart from IM operations, this scenario was typical in terms of workload for an arrival at a busy airport"	99.5	6.46	6.00
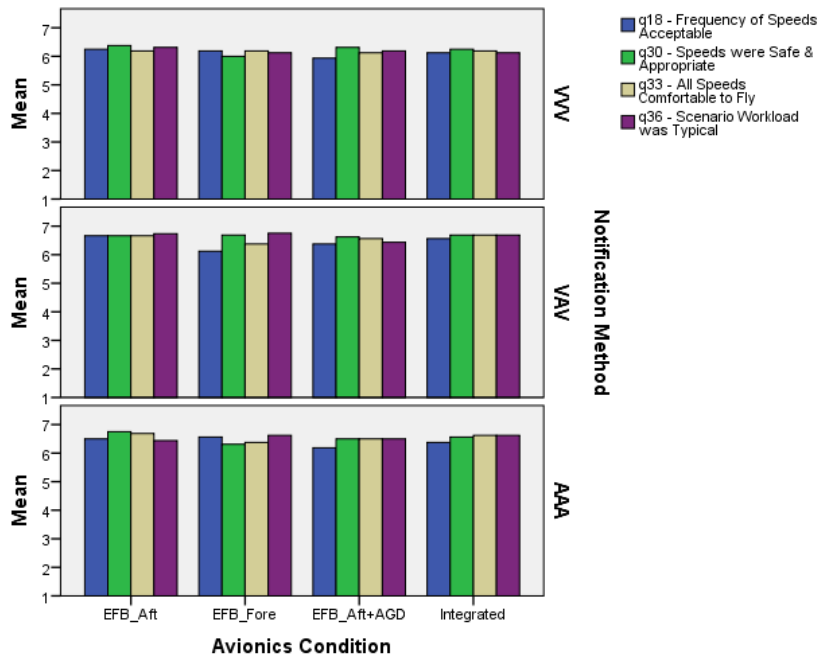
Figure 34. Mean Ratings on Scenarios (Scale ranges 1 to 7).
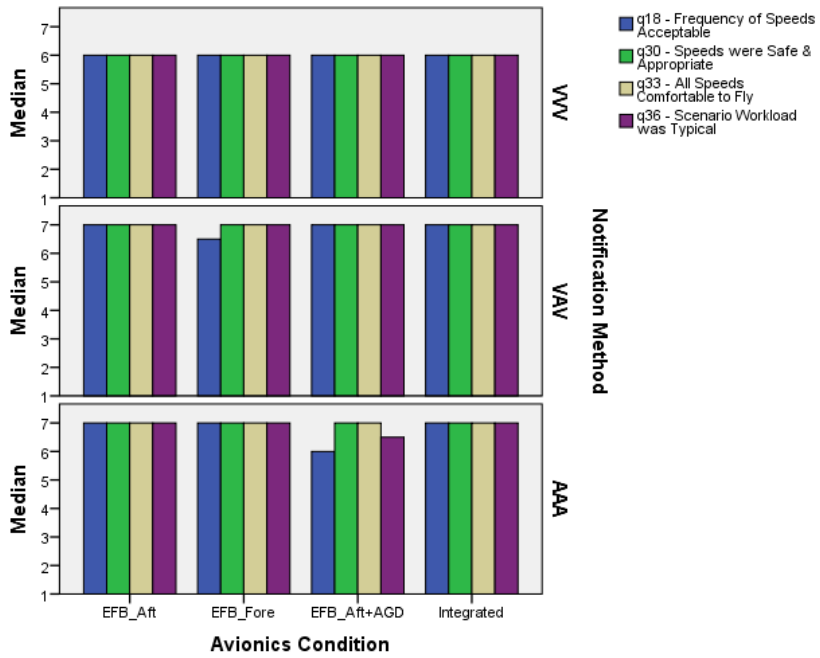
Figure 35. Median Ratings on Scenarios (Scale ranges 1 to 7).

5.3 Post-experiment Questionnaire Data

Post-experiment questionnaire items asked pilots to consider pairwise preference comparisons and to also rate (using 9-point scales with anchoring cues) the operational acceptability of the Avionics Conditions in the context of the notification method they received, the utility of aural indications, and factors associated with operational acceptability (workload, situation awareness, and crew coordination). Appendix E presents the post-experiment questionnaire items analyzed in this section.

Simulation Validity and Effectiveness Ratings

Crews were asked to provide ratings on the effectiveness of the training they received and the validity of the simulation environment and scenarios. Figures 36 and 37 show the means and medians of these ratings, respectively. All ratings were more positive than the midpoint rating (Binomial test $p<0.001$), and medians were at least 7.5 on a 9-point scale, with not score less than 6. Notification Method did not affect these ratings (all $p≥0.352$).
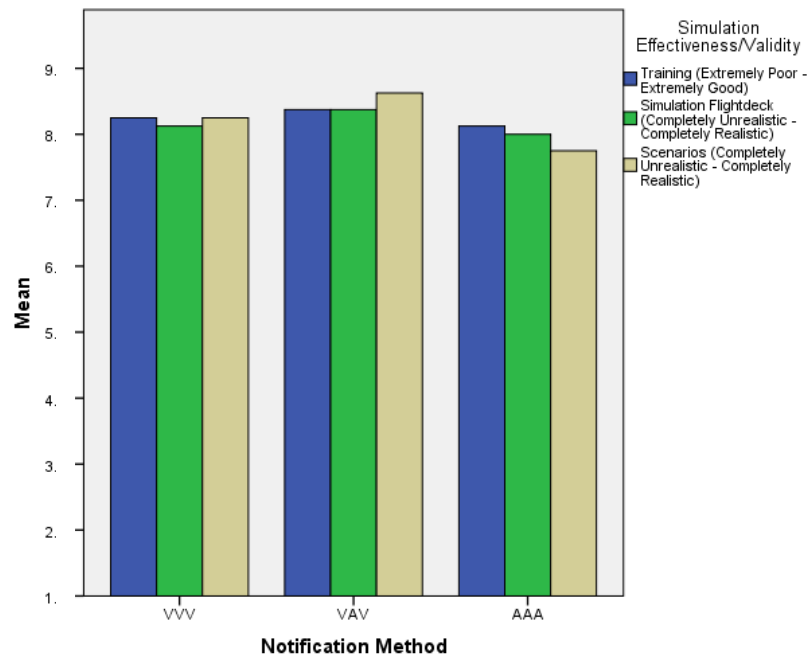
Figure 36. Mean Study Effectiveness Ratings (Scale ranges 1 to 9).
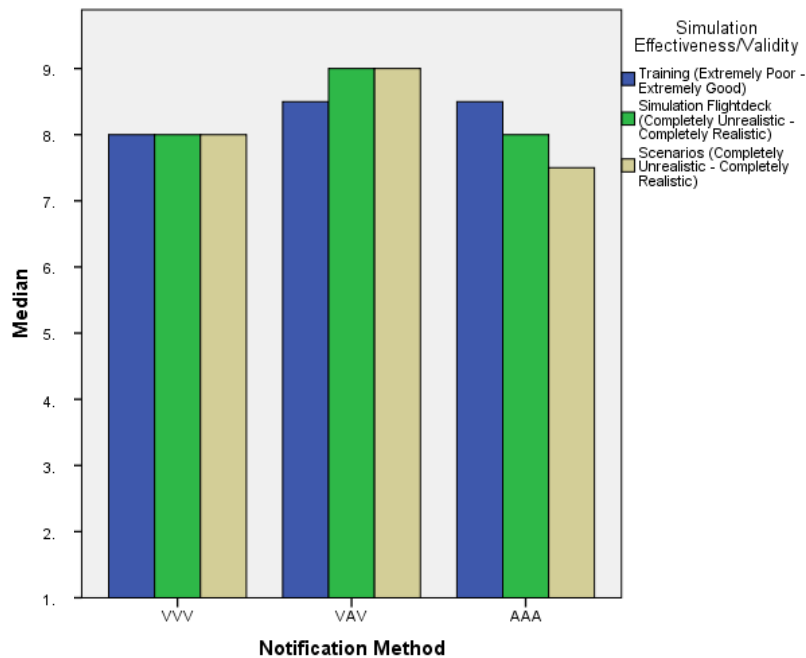
Figure 37. Median Study Effectiveness Ratings (Scale ranges 1 to 9).

Preference & Operational Acceptability Ratings

Crews were asked to provide relative preference ratings for each pair of Avionics Conditions using bipolar scales where each side of the scale indicated one of the Avionics Conditions, and the term "Strongly Prefer" at both ends of the scale and "No Difference" as the central response cue. Preference percentile ratings were derived from these data according to the Analytical Hierarchy Process (Saaty, 2008) which results in a priority rating among the rated set of alternatives (Figure 38). The Integrated condition received the highest percentage preference for all Notification Methods.

Crews were consistent in the Avionics Configuration that they preferred least and most. All but one pilot indicated the lowest preference for the EFB_Aft condition and highest preference for the Integrated condition. The one dissenting pilot preferred the Integrated condition the least and preferred the EFB_Fore condition the most. Figure 39 presents the number of pilots who selected each of the possible display conditions: the EFB_Aft+AGD and the EFB_Fore conditions. When aural notifications were used for all IM events (AAA), the EFB_Fore condition

was more likely preferred. When visual-only indications were provided for commanded speed onsets and as a reminder to implement these, the EFB_Aft+AGD was preferred.

Pilots also rated the Operational Acceptability (*i.e.*, *in terms of operational risk, heads down time, workload, total situation awareness, scan pattern, etc.*) for each Avionics Condition by selecting a value of 1-9 describing their degree of agreement (1 for "Strongly Disagree" to 9 for "Strongly Agree") with the statement, "It is acceptable to conduct FIM operations with (each type of Avionics)." Figures 40 and 41 show the means and medians of these responses.

Avionics Conditions were rated significantly different (p=0.001), whereas the Notification Method that the pilots experienced did not significantly affect ratings (p=0.741). For the Avionics Conditions, all pairwise comparisons are significantly different (all p≤0.038); and these reflect pilots' consideration of these conditions in the following order, from most acceptable to least: Integrated, EFB_Fore, EFB_Aft+AGD, EFB_Aft. Binomial tests on each rating for Avionics Condition show that significantly more of the pilots rated the EFB_Aft condition in the lower half of the scale (p=0.023), the median response could not be distinguished from the midpoint ("Borderline") of the scale for EFB_Aft+AGD ratings, and both the EFB_Fore and the Integrated condition received ratings in the better half of the scale (p=0.023, p=0.002, respectively). These results generally agree with those identified through the preference rating questionnaire items.
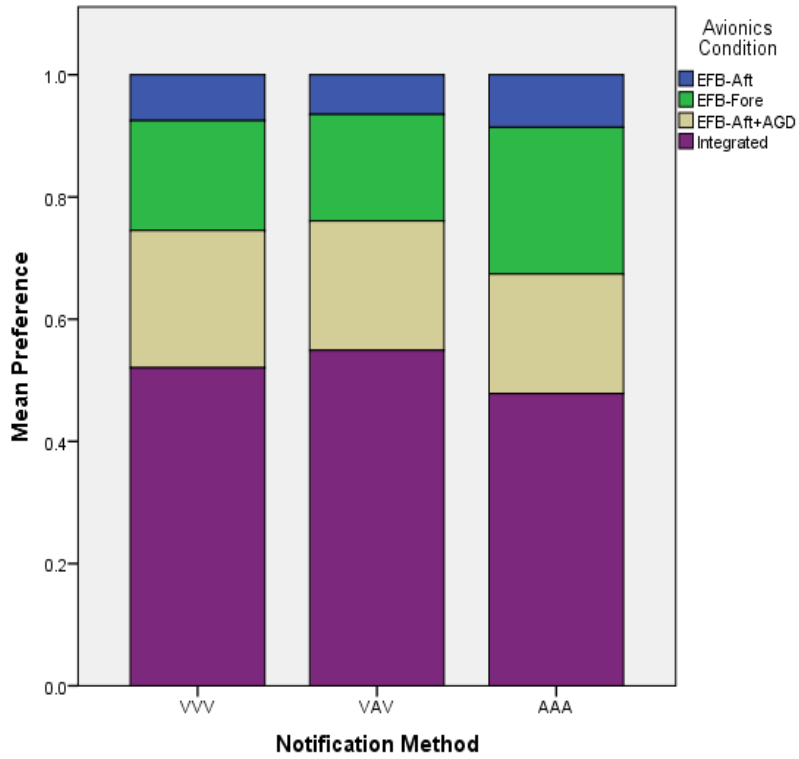
Figure 38. Apportionment of Preference (Scale ranges 0 to 1).
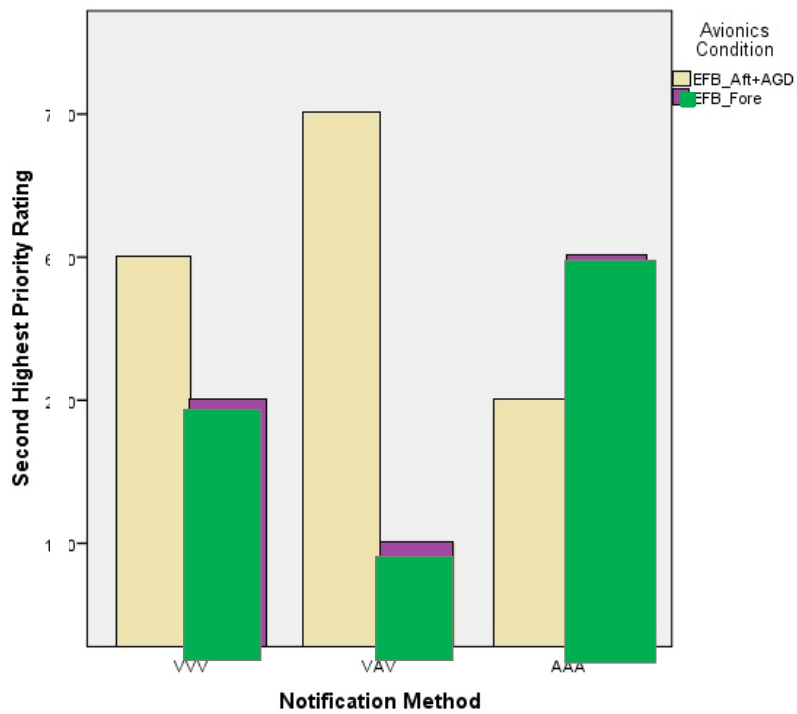
Figure 39. Counts for Second Highest Avionics Condition by Notification Method.
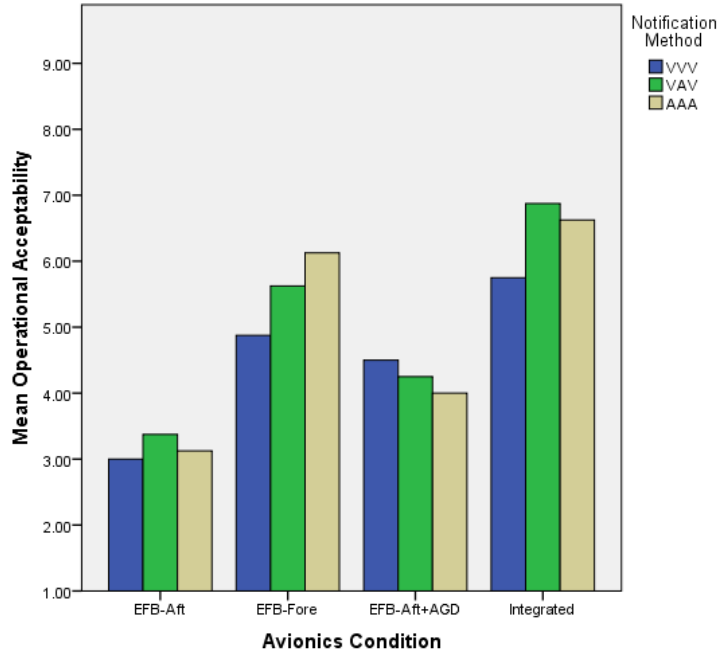
Figure 40. Mean Operational Acceptability of Avionics Conditions Ratings.
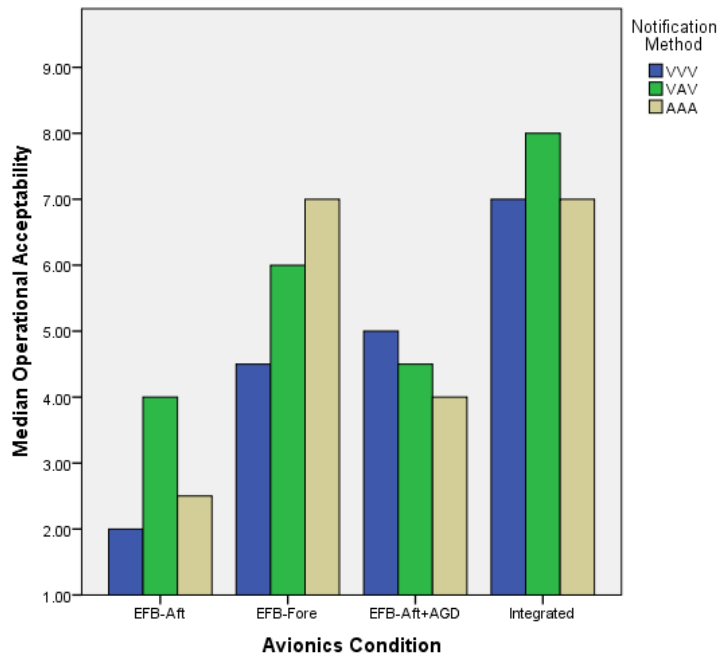
Figure 41. Median Operational Acceptability of Avionics Conditions Ratings.

Ratings on Aural Notifications

Pilots were asked to rate the aural indications when used to indicate commanded speed onset, as a reminder, and for deviations from the commanded speed. While two separate scales (unnecessary /necessary and distraction/appropriate) were provided for each of the 3 questions, several pilots interpreted the two scales as one scale and provided only one response. Where the responses were properly scored on two scales, the ratings differed in 7 out of 72 opportunities, and the pilots' ratings on the distraction scale had higher scores than the necessity scale. In one case, the necessity and appropriateness ratings differed by 2; all others differed by no more than 1 scale unit. It seemed that pilots' ratings on these scales, even when two ratings were provided, represented conflated concepts. As such, ratings on these two scales were averaged for the results presented below. Scores of 1 indicated that the aural indication would be considered unnecessary/distracting, whereas a score of 9 indicates that the aural indication was considered necessary/appropriate. Note that not all conditions actually used aural indications. The AAA method had aural indications for all three events, the VAV only for conformance deviations, and the VVV method had no aural indications. Figures 42 and 43 show the means and medians of the composite ratings.

Ratings were not statistically different by IM Event (p=0.402), however Binomial tests show that, in general, ratings of aural notifications for target speed onset and for reminders were in the more positive half of the scale (p<0.023), whereas ratings on the use of aural notifications for speed deviation indications were indistinguishable from the central scale response (p=0.152). Notification Method was a factor in how pilots rated the appropriateness of aural notifications for these IM events (p=0.006). Pairwise comparisons showed that pilots who experienced the AAA method rated the use of these aural notifications more positively than those who experienced the VAV method (p=0.029). Also, pilots rated AAA more positively than those who experienced no aural indications, the VVV method (p=0.018). There was no significant difference in the ratings for the VAV and VVV methods (p=0.648).
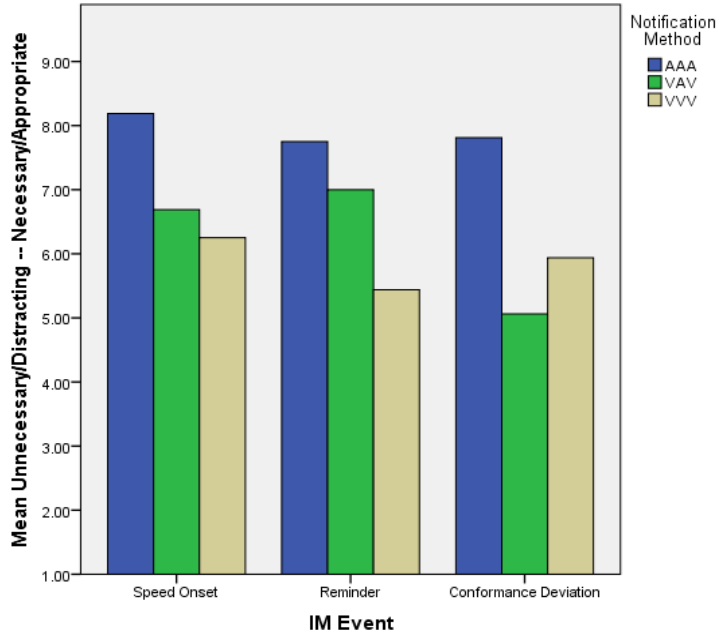
Figure 42. Mean Rating for Appropriateness of Aurals by IM Event.
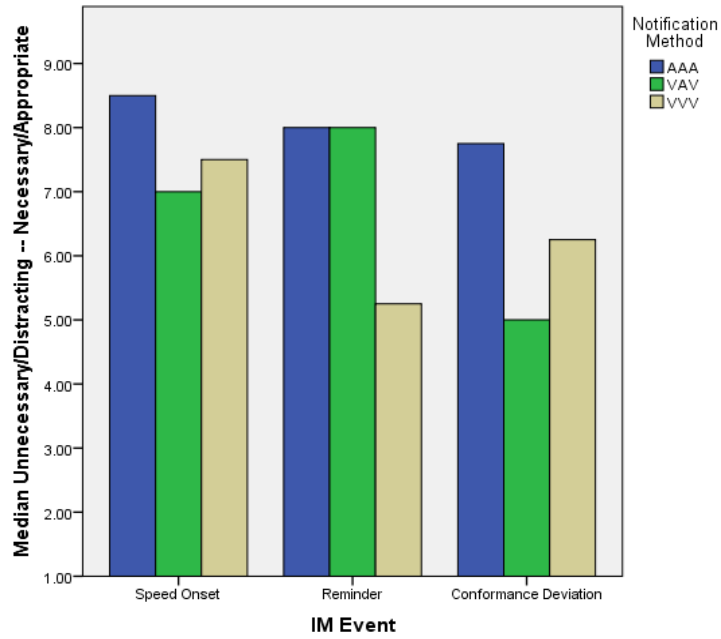
Figure 43. Median Rating for Appropriateness of Aurals by IM Event.

Reminder & Conformance Thresholds Ratings

All crews received both reminders and conformance deviation indications. However, whereas only half of the pilots recalled receiving reminder indications, all pilots indicated that they had received conformance deviation indications. Those who recalled receiving either a reminder or speed conformance deviation indicator, were asked to rate the degree to which they considered these events false alarms; where a 1 indicated "Almost Never," a 5 indicated "50/50," and a 9 indicated "Almost All." Figures 44 and 45 show the means and medians for these ratings respectively.

Binomial tests indicated that most pilots rated reminders as accurate (less than the midpoint of this scale) ($p=0.039$). However, half of the pilots provided ratings greater than the midpoint of the scale for conformance deviations– so considered these more likely to be false alarms ($p>0.999$). This significant effect is revealed as a significant main effect of IM type on these ratings ($p=0.005$). While Notification Method did not significantly affect these ratings ($p=0.823$), the interaction of this term and IM event was significant ($p=0.033$). As observed in the means and medians, the condition causing this significant effect pertains to the ratings from pilots experiencing the VAV method. Those pilots who experienced conformance deviations annunciations were more likely to interpret the annunciation as a false alarm compared to the reminders.
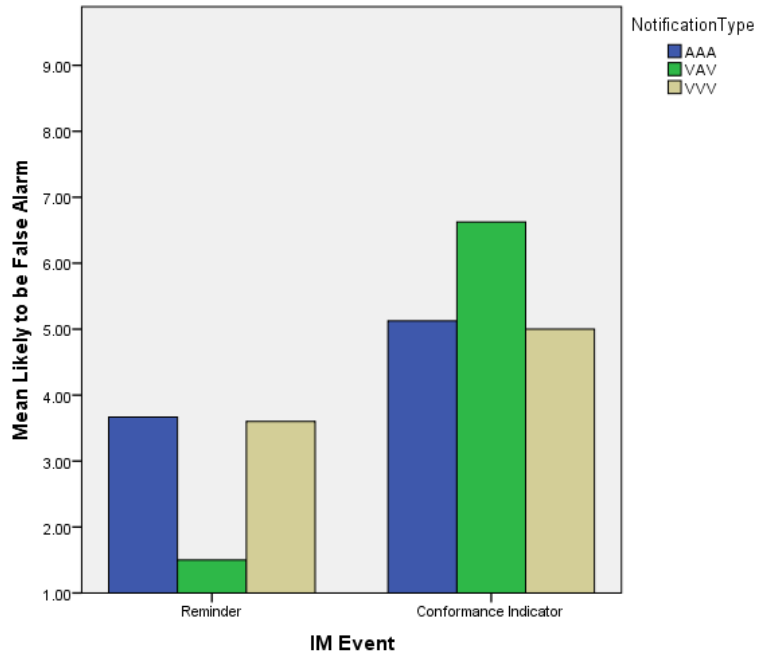
Figure 44. Mean Rating for 'Probability IM Event is a False Alarm'.
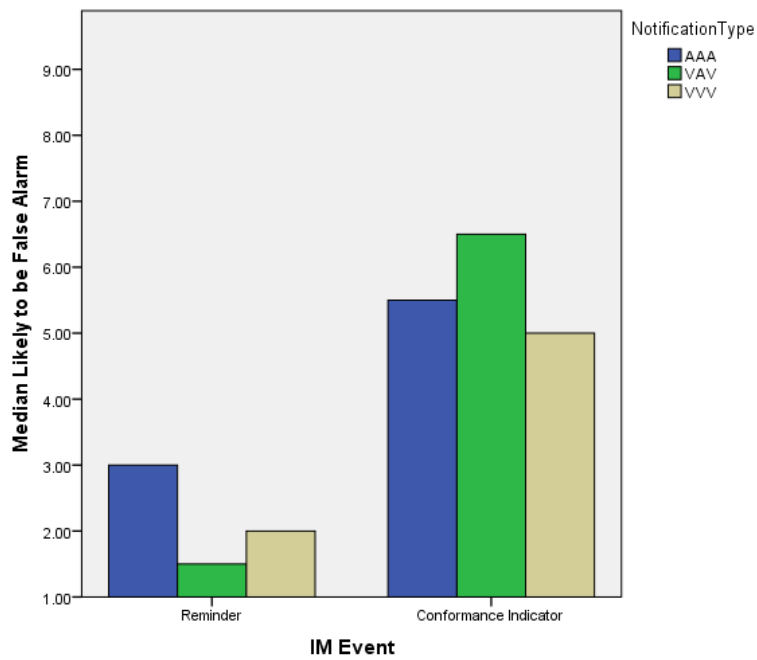
Figure 45. Median Rating for 'Probability IM Event is a False Alarm'.

Operational Impact Ratings

Crews were asked several questions regarding the operational impact of flying IM operations compared to current operations. The following questions were rated using a 9-point scale (with these anchoring terms):

1. *"How would you rate the overall effect of FIM operations on your workload, as compared to normal operations in similar aircraft under similar conditions?" (Greatly increases workload - Greatly reduces workload);*
2. *"How would you rate the overall effect of FIM operations on your situation awareness of arrival speeds, as compared to normal operations in a similar aircraft under similar conditions?" (Greatly reduces situational SA - Greatly increases SA);*
3. *"How would you rate the overall effect of FIM operations on situation awareness of other elements, as compared to normal operations in a similar aircraft under similar conditions?" (Greatly reduces SA - Greatly increases SA);*
4. *"Flying commanded speeds via the auto-throttle / MCP was acceptable" (Strongly Disagree - Strongly Agree);*
5. *"How well did your standard crew coordination and cockpit resource management strategies map onto the use of FIM operations?" (Extremely poorly, specific training required - Extremely well, no adaptation required).*

Figures 46 and 47 show the mean and median responses on these scales. Notification Method experienced by pilots did not differentially affect their ratings of the operational impact of IM operations in terms of perceived workload relative to standard operations (p=0.104), perceived situation awareness of arrival speeds (p=0.134), the acceptability of flying target speeds with autothrottle/MCP entry (p=0.166), or the degree to which standard crew coordination and cockpit resource management strategies map to IM operations (p=0.600). Notification Method that pilots experienced was a significant factor in explaining variability associated with ratings for perceived situation awareness of other elements during flight (p=0.041). However, pairwise comparisons reveal ratings that are significantly different between the AAA method and the other two methods (both p=0.091), where ratings for general SA were higher for those who received the AAA method.

Averaging over all Notification Methods, significantly more than half of the participants provided positive ratings regarding operational impact with respect to all measures: workload (p=0.064),

general situation awareness (p=0.064) and situation awareness of arrival speeds (p<0.001), method for flying speeds (p=0.023), and crew coordination / resource management (p<0.001).
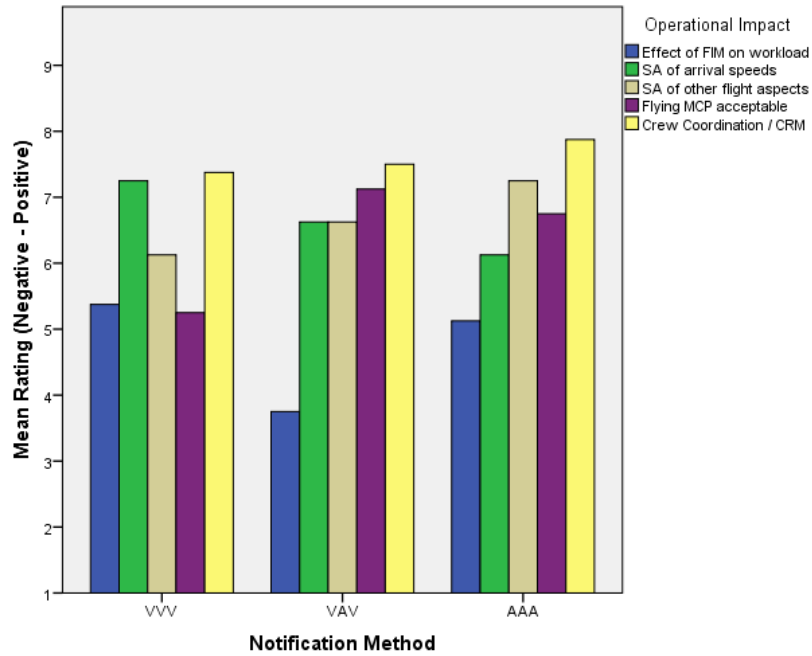
Figure 46. Mean Rating for Operational Impact Questions.
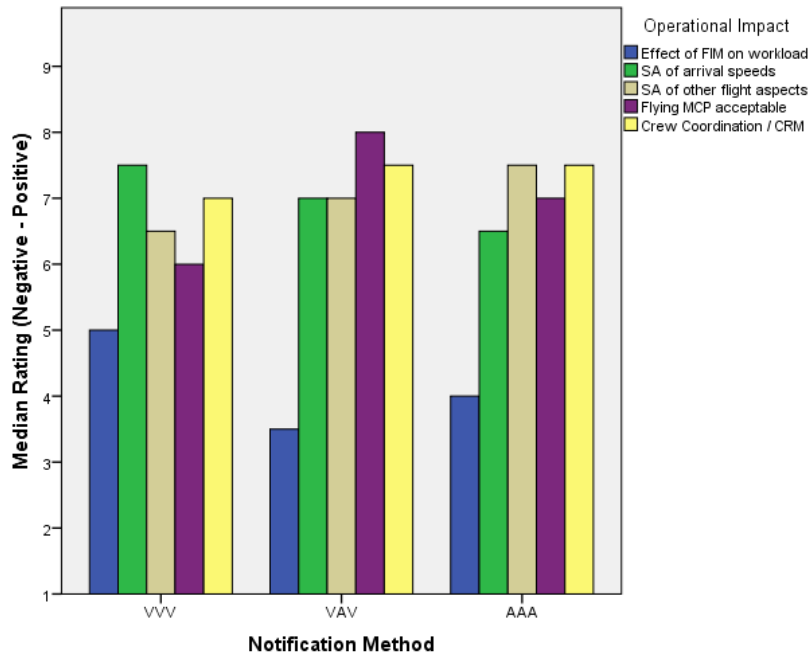
Figure 47. Median Rating for Operational Impact Questions.

Participant Comments

Pilot comments were obtained from the post-experiment questionnaire and during debriefing discussions.

Comments on Avionics Conditions

In addition to explicit questions regarding preference, pilots were asked to provide comments on their relative impressions of the different Avionics Conditions. Twenty of the pilots provided positive commentary on the Integrated condition, and 16 commented positively on the EFB_Fore condition. When pilots were critical of the Integrated condition, it was only to suggest that the font size of the target speeds was too small, and that the flashing reminders did not persist for long enough. Those critical of the EFB_Fore suggested that it was mounted in a position that was too close to comfortably read, and that there might be some concerns with its physical location in turbulent conditions. Pilots noted a concern for vertigo with this condition, and more so with the EFB_Aft condition. Pilots who commented on the AGD routinely mentioned that they overlooked changes to target speeds and reminders on this device. While it was in the primary field of view, it is at a different focal distance than instruments in the panel. In addition, the lights used to indicate new speeds were considered not salient enough to attract attention. Only two subjects indicated that the EFB_Aft condition was useable.

Comments regarding the best way to direct attention to FIM events focused on the use of aural indications. While all pilots mentioned whether they would or would not want aurals to augment visual indications, most pilots did not comment on the type of visual notifications they desired. Those pilots who did comment noted the appropriate use of flashing to indicate FIM events. Four pilots suggested that this formatting method be used for all FIM events, three pilots mentioned onset only, three pilots mentioned reminders only, and one pilot mentioned both onsets and reminders. Three pilots indicated that changing display element color might be effective for reminders and conformance, and one pilot suggested that reminders and conformance deviations be presented in red.

Comments on Use of Aurals

Pilots were asked how the three types of FIM events (new speed target, conformance deviation, reminder) should be brought to their attention. Table 2 summarizes these results.

Table 2. Appropriate use of Aural Indications by IM Event.

IM Events (Notification Method tested)	VVV (8 total)	VAV (8 total)	AAA (8 total)	Total (%)
Onset, Reminder & Conformance (AAA)	5	1	6	12 (50%)
Onset & Reminder	0	1	0	1 (4.2%)
Onset & Conformance	1	0	1	2 (8.3%)
Reminder & Conformance	0	2	0	2 (8.3%)
Onset only	2	1	0	3 (12.5%)
Reminder only	0	2	0	2 (8.3%)
Conformance only (VAV)	0	0	0	0 (0%)
Visual only (VVV)	0	1	1	2 (8.3%)

Comments on Alerting Thresholds

Pilots were asked to comment on the logic associated with receiving reminders and conformance deviation indications. Most (11 pilots) of those who commented on the reminder logic (14 pilots) agreed with the implementation tested, notifying the crew if the speed target had not been dialed into the MCP within 10 seconds. Two pilots suggested slightly longer notification times, while one suggested slightly shorter. The conformance logic involved three parameters: the deviation of actual speed from the intended speed profile, the time since the speed change, and whether the speed deviation is trending towards convergence on the profile speed or divergence from it. Crews did not always express their preferences in terms of all these parameters. Six pilots suggested conformance indication logic strictly in terms of a speed deviation from the profile speed, four of these indicated a 7 knot deviation, and two indicated a 10 knot deviation. Four pilots indicated that the logic should be based on the time that the aircraft is not on the correct profile speed; two suggested 15 seconds, one suggested a range of 10-20 seconds, and another a range of 15-20 seconds. Eight pilots expressed logic in terms of both a speed deviation and an elapsed time: 5 knots for 10 seconds (1 pilot), 7 knots for 10 seconds (2 pilots), 10 knots for 10 seconds (1 pilot), 10 knots for 15 seconds (1 pilot), and 7 knots for 20-30 seconds. Two of pilots mentioned that the degree to which the aircraft is trending towards the profile should be considered, and two indicated that the logic should factor in the magnitude of the speed change required. One pilot mentioned that the logic should

trigger a notification only if the deviation impacts operations significantly, i.e., achieve-by point or traffic.

Comments on Operational Impact

The pilots' comments regarding workload, situation awareness, and crew coordination were generally positive. Pilots were concerned with distraction from scanning outside, the potential for obtaining speeds that are difficult to achieve, and the impact on passenger comfort and fuel use from increased speed brake use compared to current operations. While pilots generally reported that standard crew coordination and resource management techniques served them well in conducting FIM operations in the simulator, several pilots did mention the need to emphasize callouts when a new target speed is received and when speed deviations are detected.

5.4 Oculometer Data

Oculometer data helps characterize the attentional sampling pilots used in response to the different Avionics Configurations presenting FIM information. The selection of data appropriate for consideration of each measure is presented per section, as these varied.

Sampling the FIM Display

FIM Display Sampling analyses were conducted on the portion of each scenario from 99 seconds into the run, until 19 seconds following the eighth speed target encountered. This period was determined by attempting to maximize data used, minimize disproportionate lost data, include an equivalent number of speed target changes, and attempt to have roughly equivalent task durations. While the number of data points taken in these windows differed only slightly, counts were normalized by data frames per scenario.

Results show that only the Avionics Condition significantly predicted differences in the counts of POGs on FIM displays ($p < 0.001$). All pairwise comparisons of Avionics Conditions show significant differences (all $p < 0.032$). Over the course of the runs, pilots were most likely to sample the Integrated FIM Display; of the retrofit conditions, pilots were more likely to sample

71

the FIM Display(s) associated with the EFB_Fore, then EFB_Aft+AGD, and least likely to sample the FIM Display associated with the EFB_Aft condition (Figure 48).
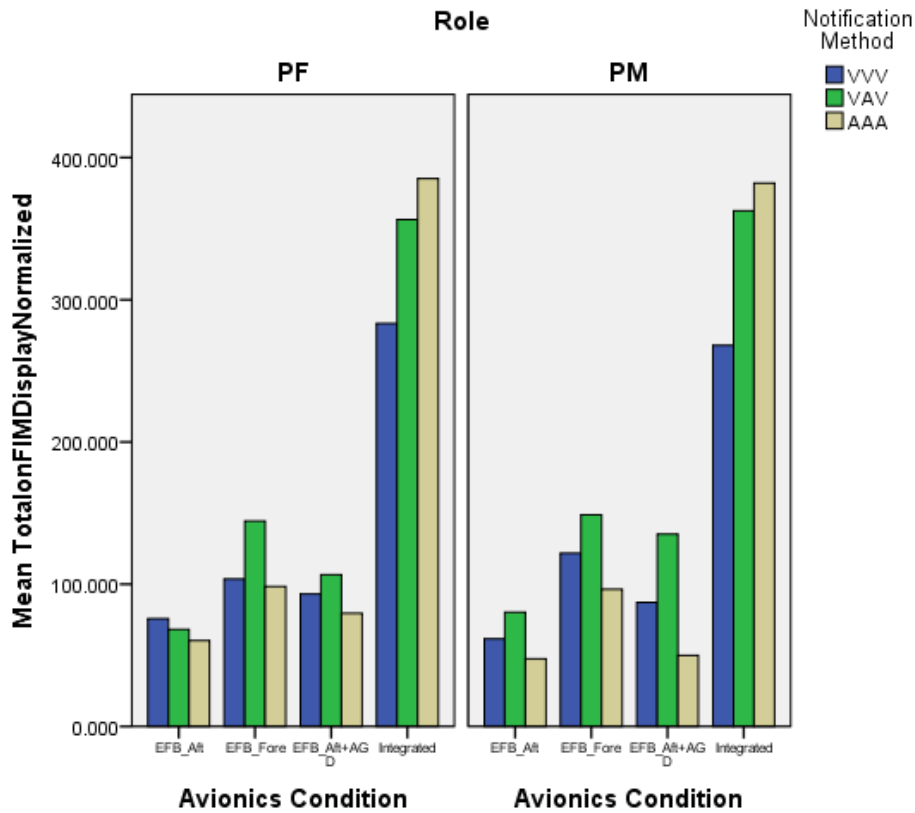
Figure 48. FIM Display Sampling.

Heads-Up Sampling

The analysis presented in this section assesses whether experimental conditions differentially affected POGs associated with Heads-Up gazes. The set of data used in this assessment was defined in the same manner as for the FIM Display Sampling analysis. However, whereas the FIM Display analysis used only data in which gaze direction quality was sufficient, this analysis employs a technique developed at NASA Langley (Ellis, Arthur, Latorella, Kramer, Shelton, Norman & Prinzel, 2012) to define Heads-Up gazes from head pitch data when gaze quality is questionable.

Means of normalized Heads-Up POGs are shown in Figure 49. Avionics Conditions significantly affected Heads-Up POGs (p=0.005). Pairwise tests show only one significant comparison; this indicating that pilots experienced significantly more Heads-Up POGs in the Integrated Condition than for the EFB_Aft Condition (p=0.016), where all other comparisons were not significant (all p>0.106). The Avionics Condition and Notification Method interaction term was significant (p=0.011), but no pairwise comparisons reached significance (all p>0.438).
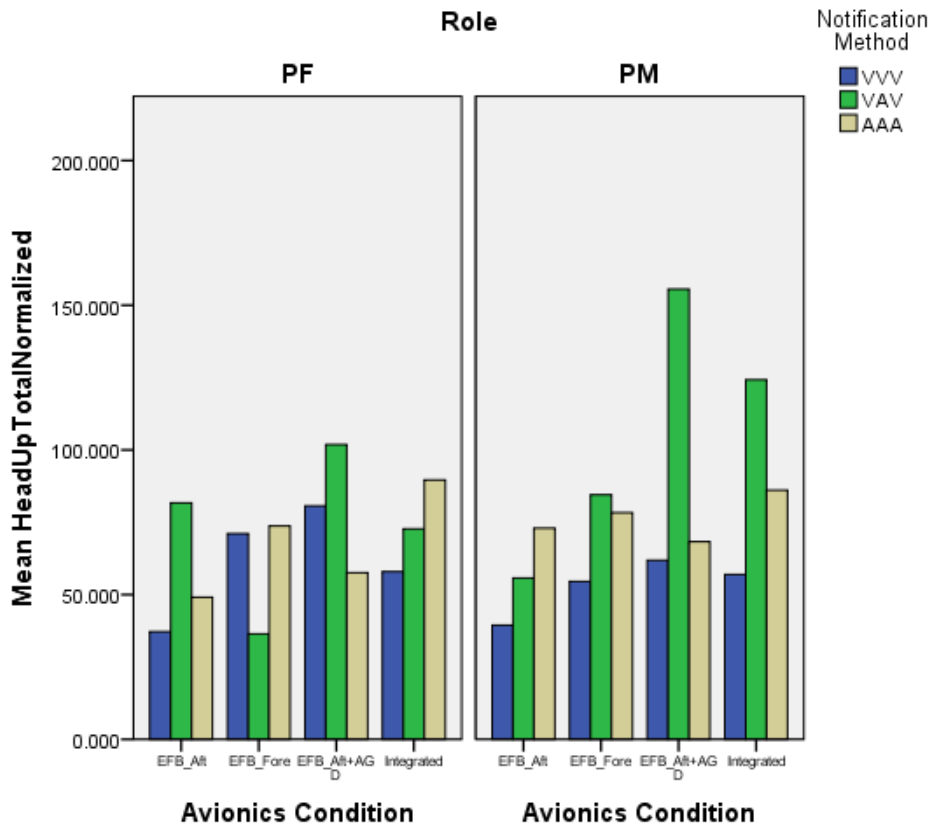
Figure 49. Heads-Up Sampling over Scenarios.

Noticing Times

Reaction times to dialing in new speed targets were identified as the interval from the time the new speed target was presented, to the time that the speed entered in the MCP's speed window changed. While this is important to understand when this aspect of the FIM operation is actuated (and these are analyzed in the main report), analysis of those data do not precisely answer the question as to how effectively particular Avionics Condition/Notification Methods combinations exogenously direct pilots' attention. The analysis presented here addresses the

latter question by identifying the Noticing Time for *both* the PF and PM to regard the display (AOI) containing information about FIM speeds. Appropriate AOIs are defined by Avionics Condition. For the EFB_Aft and EFB_Fore conditions, the appropriate AOI for reporting an appropriate POG is the EFB (where the locations of the AOI change with the location of the EFB). For the EFB_Aft+AGD condition, the first Noticing Time was identified as a POG on either the EFB or the AGD. For the Integrated condition, the PFD is the relevant AOI. Noticing times were identified in logfile data, and conducted on periods following each of eight speed targets per run. These were defined as the first POG that landed on the appropriate AOI for which the data quality was 0.7 or greater, and for which there was a second such subsequent gaze, with fewer than five frames (88msec of data) of intervening missing or poor quality data in the same AOI.

Noticing Times were significantly affected by Avionics Conditions (p < 0.001), the interaction of Avionics Condition and Notification Method (p=0.001), and Role (p=0.077). Noticing Times, and the variability in these, seemed to decrease across conditions in this order: EFB_Aft, EFB_Fore, EFB_Aft+AGD, Integrated (Figure 50).   Noticing time with the Integrated condition is, on average, over five times faster than with the EFB_Aft condition.   Pairwise comparisons show the Integrated condition to be significantly faster than all other conditions, and the EFB_Aft+AGD condition to be significantly faster than both the EFB_Fore and the EFB_Aft conditions (all p< 0.013). This main effect contains a significant interaction of the Avionics Condition and the Notification Method which shows that, for the EFB_Fore condition, Noticing time for the AAA condition is significantly longer than for the VAV condition (p=0.084). For the other three Avionics Conditions, pairwise comparisons of Notification Methods did not significantly differ (all p≥0.250), but means suggest that pilots with the VVV method were slowest to notice new speed targets. PFs were faster to notice commanded speed changes than PMs, by about 200msec.

 The former analysis looks at Noticing times for each crew member separately. The same data set was similarly analyzed to investigate how crews' Minimum Noticing Time, and the difference between pilots' Noticing Times were affected by experimental conditions. Factors of significance for these variables are similar to findings observed for each pilot's Noticing Times. Avionics Condition (p<0.001), and the interaction of Avionics Condition and Notification Method (p<0.001) significantly affected the crews' first notice of a new speed target (Figure 51).

With regard to the main effect, means followed the same order as for Noticing Times, but were more sensitive to differences in conditions. Pairwise comparisons showed only significantly faster Minimum Noticing Times for the Integrated condition than other conditions (all $p < 0.065$). Pairwise tests of interaction terms show that the AAA Method was associated with significantly faster Minimum Noticing Times than the VAV Method (with the EFB_Aft Condition, $p = 0.086$) and the VVV Method (with the Integrated Condition, $p = 0.006$).

For the same data periods used to assess pilot and crew Noticing Times above, when a Notice was detected, a count was kept for how many times the PM was the pilot to notice first. These counts were normalized by the number of Noticing Times indicated and are shown in Figure 52. Analysis as a Poisson distribution with a log link function shows only a significant effect of Notification Method, whereby PMs are more likely to be first to notice new commanded speeds than PFs when using the VVV method than the AAA method
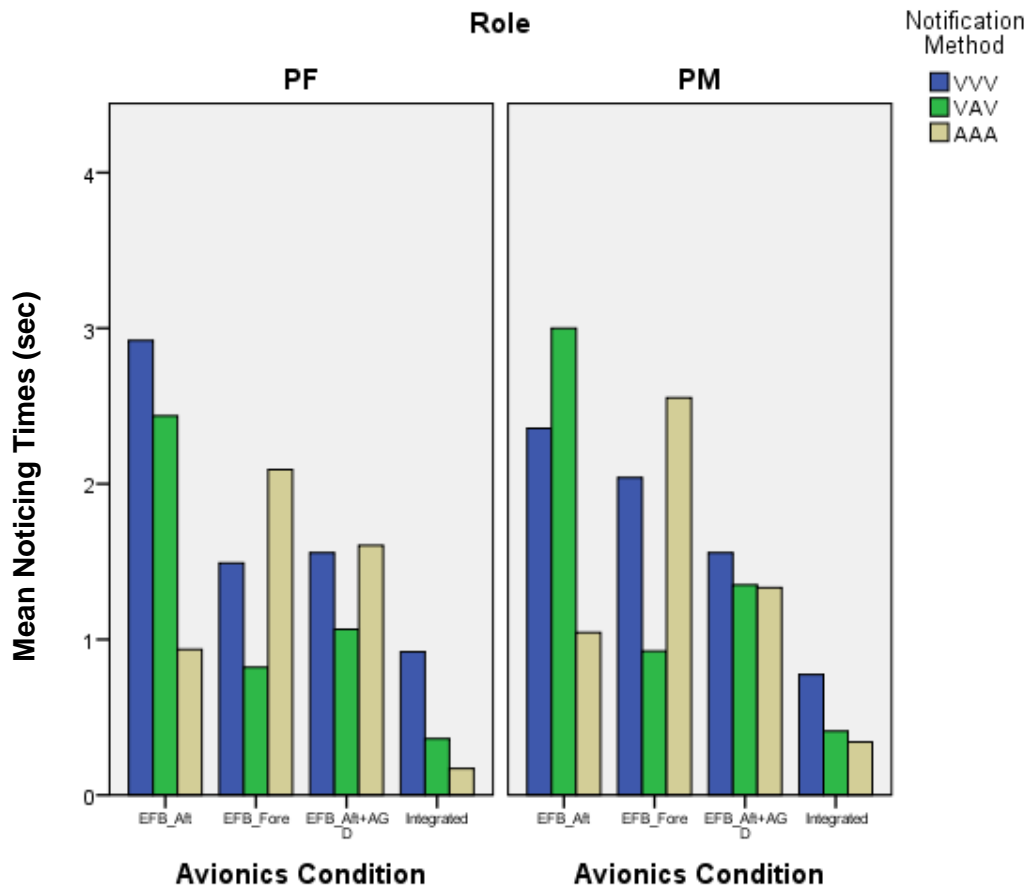
Figure 50. Mean NoticingTimes (sec).

75
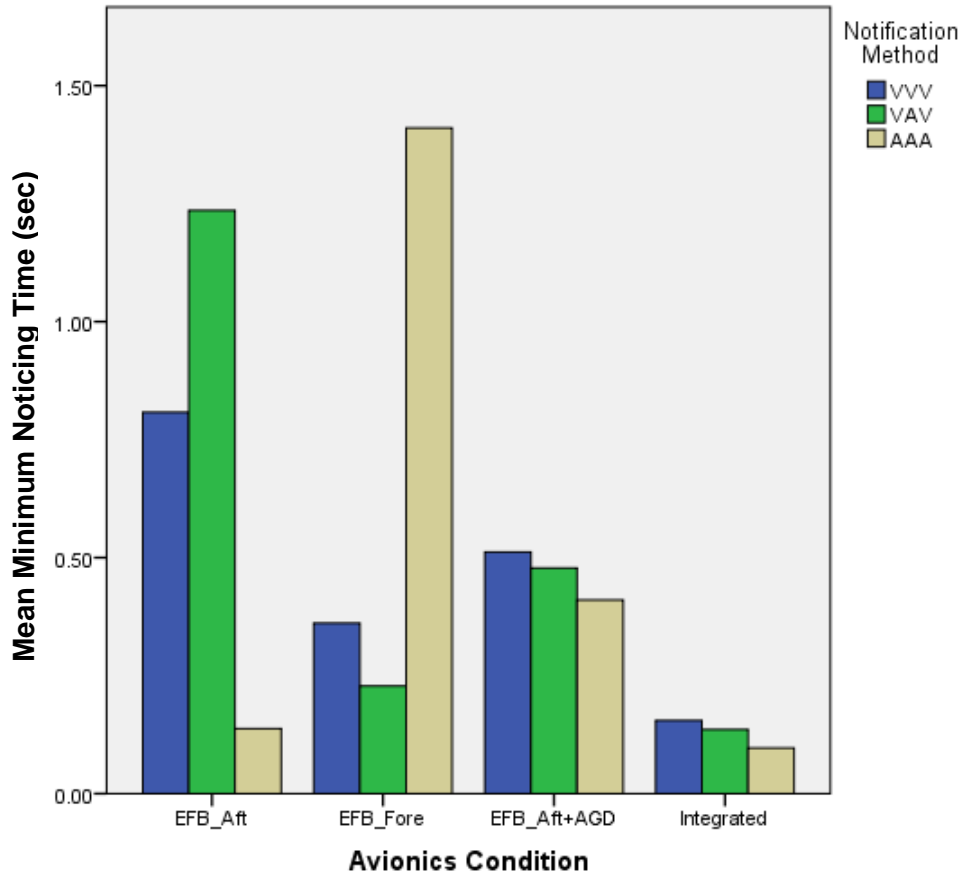
Figure 51. Minimum Noticing Times (msec).

.
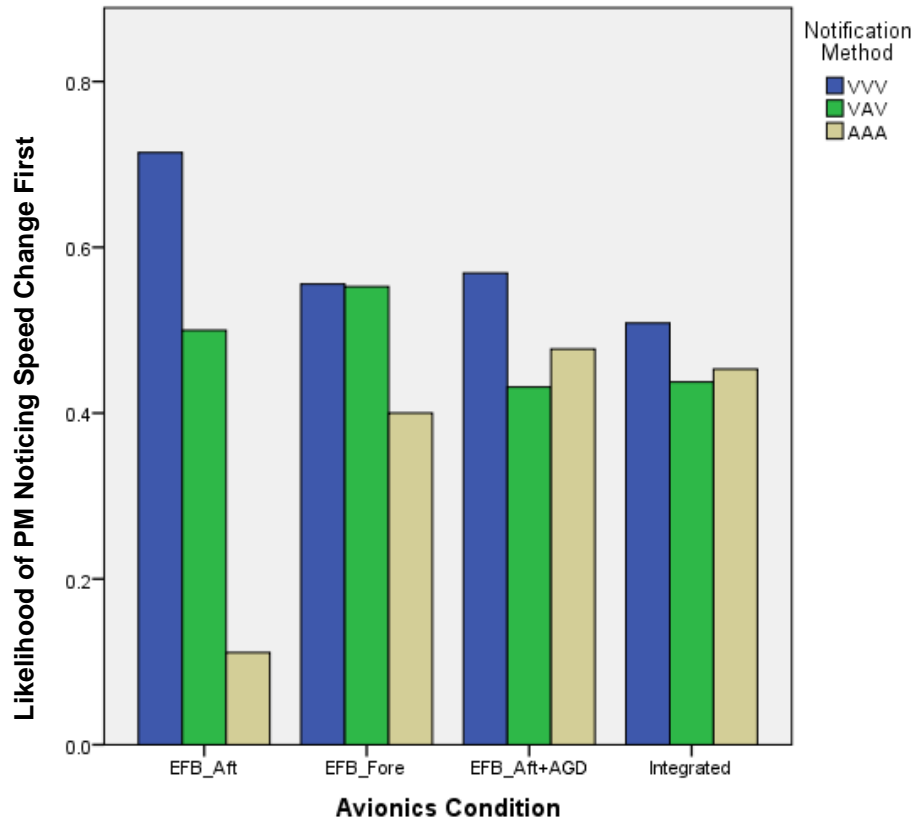
Figure 52. Likelihood of PM Being First to Notice Speed Change.

Indicators of Pilot Workload

Two measures postulated to reflect workload are examined: Dwell Time and a measure related to scan path entropy. Dwell time has related to pilots' workload (Callan, 1998), and has been thought to reflect the time required to actively process visual stimuli (Becker, 2011). In theory, more information dense, confusing, or unintuitive presentations should require longer dwell times to detect and extract pertinent information.

Shannon (1948) originally extrapolated the concept of entropy from thermodynamics to characterize the completeness of information exchange. Tole, Stephens, Vivaudou, Harris, & Ephrath (1982) applied this measure to characterize distribution of POGs; and Harris, Glover, & Spady (1986) found that as pilots' workload increased, visual sampling became more systematic and entropy decreased. Di Nocera (2007) and colleagues showed the Nearest-Neighbor Index (NNI) measure of entropy to be consistent with both objective (p300 EEG

responses) and subjective (NASA-TLX score) measures of workload (Camilli, Terenzi, & Di Nocera, 2007). The NNI metric investigated here, is the ratio of the average observed minimum distances among POGs, and the mean distance expected if the distribution were random. The NNI here is equal to one when the distribution is completely random, and higher values suggest more systematic search – presumably induced by higher workload conditions. In contrast to Spady's (1986) work in an aviation context, with fixed information channels, Di Nocera (2007) found an inverse relationship between an entropy measure and task loading for a free-viewing task.

Dwell Time on First Looks

Dwell time is a measure associated with the difficulty of extracting usable information from a gaze. Once a POG satisfied the conditions to identify that the pilot noticed the speed (see Noticing Time definition above), Dwell Time was calculated from the data frame of first Notice (POG on the appropriate AOI) until either the frame before another AOI was reported, or five frames of missing or poor quality data occurred; then this number of frames was multiplied by the nominal frame rate, 17msec/frame. Data used for this study included those for the eight speed targets encountered starting with the second of these, as for Noticing Times, again using logfile data.

Dwell times were significantly affected by the Avionics Condition factor ($p < 0.001$) (Figure 53). Pairwise comparisons of these levels show essentially a bifurcation of these conditions; the Integrated and EFB_Aft+AGD conditions did not significantly differ from each other, but they both supported significantly shorter Dwell Times than either the EFB_Fore or EFB_Aft conditions (and these last two did not significantly differ from each other) (all $p < 0.002$).
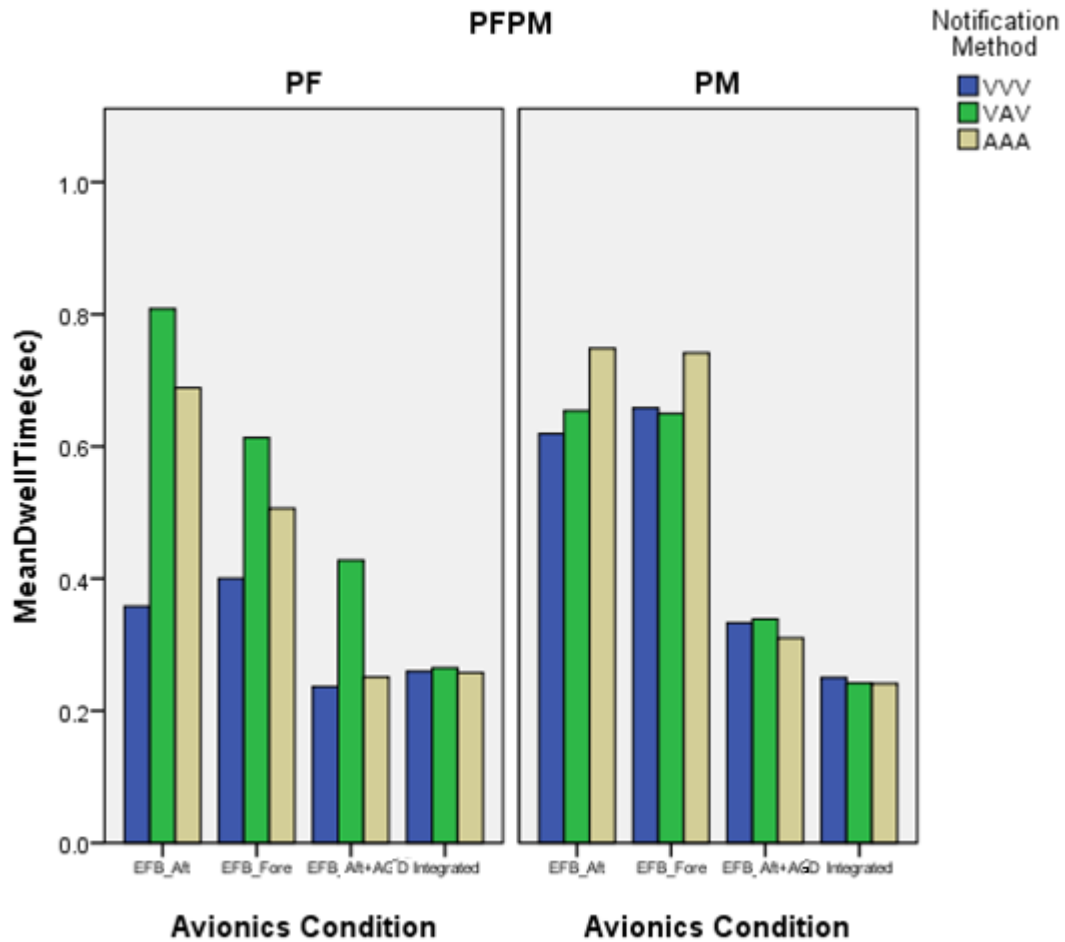
Figure 53. Mean Dwell Times (sec).

Nearest-Neighbor Index - Indicator of Pilot Workload

Nearest-Neighbor indices of Total Entropy were calculated for good quality data following the occurrence of a new speed target, and for the following 19 seconds. Data included in this analysis was from the eight speed targets beginning with the second speed target occurring in logfiles for each run. Total Entropy measures were calculated from software developed for this purpose at NASA Langley (Harden & Latorella, 2014), based on Di Nocera's publications. The 5 Hz oculometer data was used for this analysis due to processing complexity. In this implementation, Total Entropy is higher with less random scan paths.

Results indicate significant effects of Notification Method (p=0.099) and Role (p=0.024). Pairwise comparisons were not significant for levels of Notification Method (all p≥0.132).

Observation of means shows clearly higher entropy (higher workload) when pilots had VVV notifications than other Notification Methods. The significant main effect of Role appears to be due to significantly higher entropy (higher workload) for pilots when in the PM role. Figure 54 illustrates the effects of Avionics Condition, Notification Method and Role on the Total Entropy index.
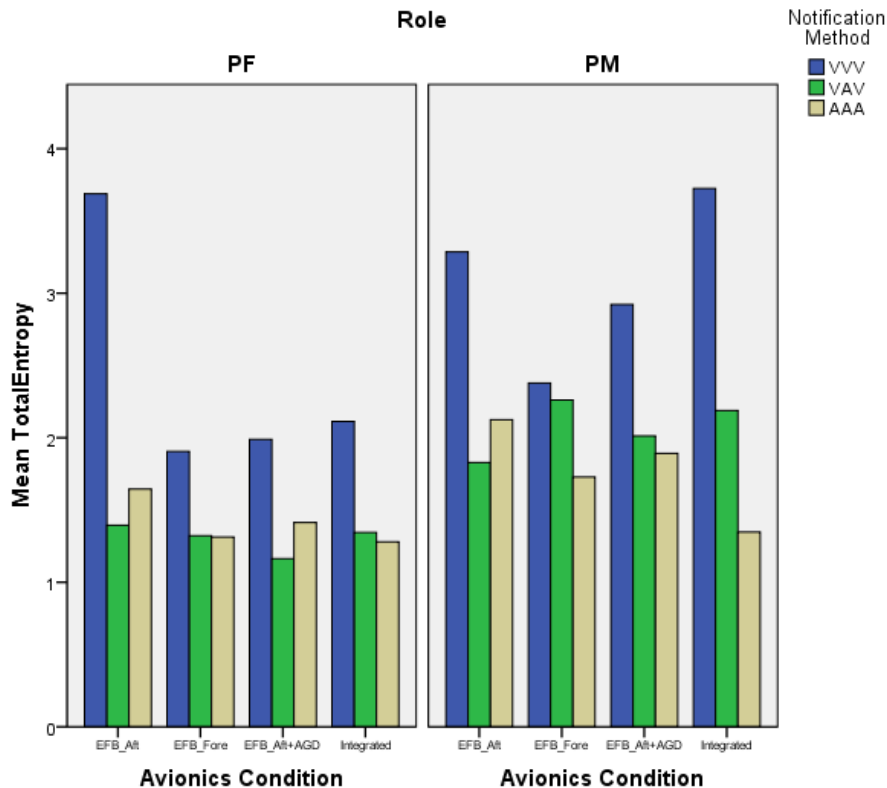
Figure 54. Mean Total Entropy Index.

Scan Path Comparisons – the Kullback-Leibler Distance Metric

Oculometer data is often used to assess differences between experts and novices, and the results of these assessments used to develop training practices. As such, development of methods for comparing attention distribution and scan paths is itself a vibrant research area. This study used the Kullback-Leibler (KLD) metric (Kullback & Leibler, 1951) to assess scan path cohesion between the PF and PM in a crew. Essentially, the KLD metric indicates the number of bits required to predict one scan distribution based on the other; so, a higher number

of bits indicates a greater difference between the two paths (Tatler, Baddeley & Gilchrist, 2005). Riche, Duvinage, Mancas, Gosselin & Dutoit (2013) indicated that the KLD is one of three metrics that is particularly informative for assessing scan path cohesion.

The KLD measure of scan cohesion was calculated for each crew's oculometer data only when both the PF and PM data was of good quality (gaze quality measure of 0.7 or better).   Data included in this assessment was the same set as included in the NNI analysis. The KLD measure was calculated from software developed for this purpose at NASA Langley (Harden & Latorella, 2014), based on Kullback & Leibler (1951) and Dewhurst, Nyström, Jarodzka, Foulsham, Johansson, & Holmqvist (2012). Here again, 5 Hz oculometer data were used for this analysis due to processing complexity.

Results indicate that only the Avionics Condition and Notification Method interaction significantly affected the KLD measure (p=0.041). Pairwise comparisons do not reveal significant differences among any of these levels (all p<0.01). Inspection of means (Figure 55) shows only a possible latent effect of Notification Method, where means for the VVV condition were consistently lower than other Notification Methods for all Avionics Conditions, indicating greater similarity of PF and PM scans.
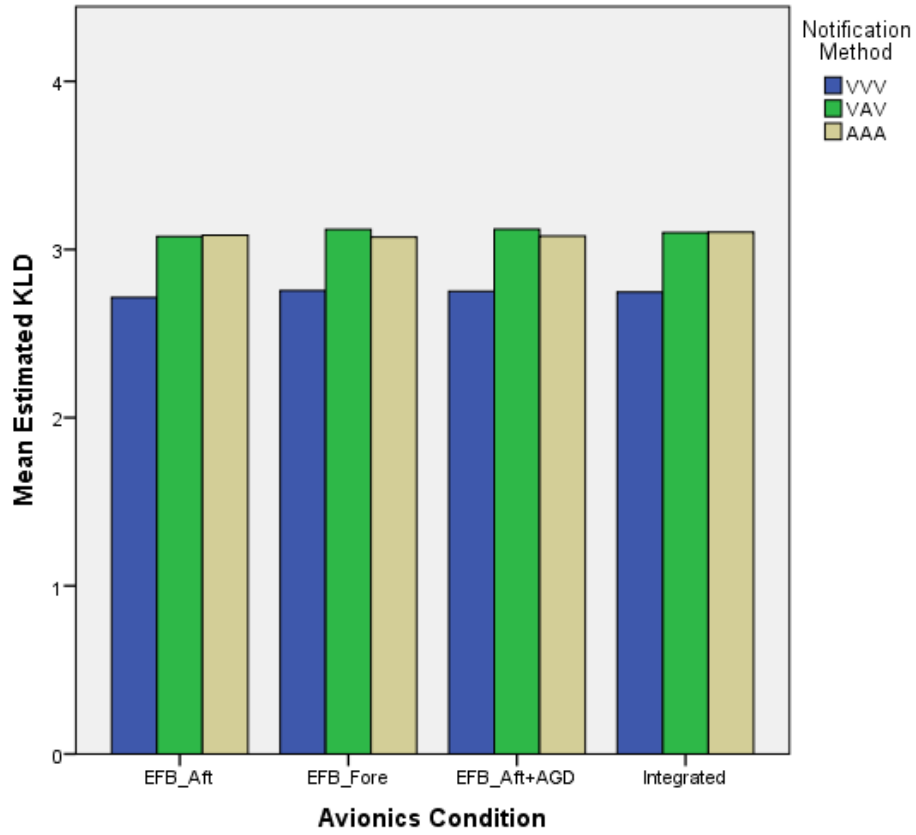
Figure 55. Mean Kullback-Leibner Divergence (KLD) Metric.

6 Summary of Results

This section summarizes the results of analyses on objective performance data, participant ratings and comments, and oculometer data for: operational acceptability of FIM operations, effects of Avionics Condition (Integrated, EFB_Fore, EFB_Aft+AGD, EFB_Aft), Notification Method (VVV, VAV, AAA), significant interactions of Avionics Condition and Notification Method, and the Role of the pilot providing data (Pilot Flying, Pilot Monitoring).

6.1 Operational Acceptability of Flying FIM Speeds

For this study, crews' IM operations included the following tasks: receiving and understanding the FIM clearance; and detecting and appropriately responding to FIM speeds from approximately top-of- descent until the final approach fix. It is important to consider that pilots who provided data regarding operational acceptability were assessing this based on a limited aspect of the IM concept of operations. In these scenarios, ATC provided the FIM clearance to the crew, but it was already entered into the FIM equipment. In addition, these scenarios offered no intentional off-nominal conditions for which pilots would, or should, consider suspending or terminating FIM operations. For this definition of IM operations, pilots provided ratings associated with operational acceptability after each run, and also in the post-experiment questionnaire.

Questions regarding the acceptability of scenarios would have been expected to be independent of the Avionics Configuration experienced. Some of these questions were unaffected by Avionics Configuration – and suggested that independent of these, pilots generally felt that commanded speeds were operationally acceptable and appropriate; the timing of these were consistent with expectations, they occurred with equal likelihood during critical tasks, and were not in conflict with other speed information; and scenario events were realistic. However, Notification Method affected three scenario questions in a similar manner (the VAV method receiving better ratings than the VVV method), and these pertained to pilots' comfort flying commanded speeds, subjective workload apart from IM operations, and (when using the EFB_Fore condition) pilots' assessment that commanded speeds were safe to fly. While overall these scenario acceptability questions received positive ratings, the significant impact of Notification Method on several indicates that this manipulation colored subjective impressions for these ratings.

In post-experiment ratings and commentary, pilots were generally enthusiastic about the concept of IM, and ratings were generally positive on scales for workload and situation awareness. Positive responses were obtained on questions that asked about the acceptability of flying commanded speeds in the MCP, the use of speedbrakes, and regarding the use of existing crew coordination strategies for conducting IM operations. Crews reported increased situation awareness with regard to both arrival speeds and other aspects of flight, compared to typical operations. While it is difficult to ascertain whether this is in fact true, it nonetheless reflects a feeling on the part of the participants that conducting IM operations is not detrimental to good situation awareness. Pilots were in general agreement that, apart from IM, scenarios presented workload that was typical for arrivals at a busy airport; and scores associated with NASA-TLX (with subscales addressing mental demand, perceived effort, perceived performance, frustration and temporal demand) and the MCH were predominantly in the low ends of these workload scales.

6.2 Avionics Conditions

Some questionnaire items were provided to participants to permit an assessment of study validity and ensure that scenarios didn't appreciably differ in ways unexpected by investigators. It was expected, then, that these measures would be insensitive to differences in the Avionics Conditions; and for many of these this was true. Responses to such questions (i.e., the acceptability of speeds provided (the speeds and timing of these, level of comfort with flying these, the frequency of these, and whether these were safe, appropriate and consistent with expectations), whether workload was typical for such an arrival, and comfort level with the degree of speedbrakes required) were in fact found to not differ among Avionics Conditions. However, other measures that might have been, were also found to not be sensitive to this manipulation. Neither a unitary questionnaire item for frustration, nor the NASA TLX frustration item were significant; nor were any of the NASA TLX subscales or the total index of workload. Ratings from the Modified Cooper-Harper workload scale, for both peak workload and average workload, were also undistinguishing; as were oculometer characterizations of workload (the NNI metric) and coherence of PF and PM scanning patterns (the KLD metric).

Where significant differences were apparent, they most apparently indicated the relative advantage of, and preference for, the Integrated condition; compared to the relative dissatisfaction and performance decrements associated with the EFB_Aft condition. This

finding is supported by converging evidence from a variety of measures. When crews used the Integrated condition, compared to the EFB_Aft condition, crews had smaller speed deviations; faster times to notice new speeds and to enter these in the MCP); higher self-ratings of general situation awareness (SART scores); higher self-reported awareness of FIM-related information (new speed commands and deviations from the speed profile); increased propensity to sample the FIM-related display, but without sacrifice to sampling the external scene (as heads-up counts were higher as well) and lower ratings for the degree of distraction induced. Because the PFD is the FIM display in the Integrated condition, it is likely that it receives sampling for its other information content. While this bias in the data must be recognized, this is also precisely a purported advantage of this implementation – it provides FIM information in a location that is already incorporated in the scan, and a main focus of attention (e.g., Diez, Boehm-Davis, Holt, Pinney, Hansberger & Schoppek (2001). This accounts for the over two-fold more likely for the FIM display for this condition to be sampled than any other condition. While it would be expected that this result would be evident for the PF, it is also true for the PM whose role is less dependent on information in the PFD. Pilots' dwell times on the first look to the FIM display after receiving a new commanded speed were faster for the Integrated condition than the EFB_Aft condition, an indication of the effort to acquire the information they desired to sample. Finally, post-experiment ratings revealed a clear preference for the Integrated condition and high scores for "operational acceptability," whereas the EFB_Aft condition was least preferred and scored significantly below the midpoint of the operational acceptability scale. Pilots' comments were consistent with these data.

It may seem obvious that the EFB_Aft condition was less desirable than other Avionics Conditions, but it was included in this test as it is the location for EFBs in some aircraft, for those would be the most expedient retrofit implementation of IM. Only one finding contradicts the conclusion that the Integrated condition offered pilots superior support for IM operations. The Integrated condition was associated with the highest incidence of speed profile deviations; particularly when paired with the VVV Notification Method. Pilots' comments regarding the Integrated condition were overwhelmingly positive, but did include mention that speed profile deviation indicator was not particularly salient. The Integrated and EFB_Aft conditions were found to significantly interact with Notification Method manipulations more than the other two Avionics Conditions. Whenever this was true, for both these conditions, the VVV method was shown to be, or considered, less supportive; or the AAA was shown to be, or considered more supportive of FIM operations. The following section explores such interactions more fully.

Given that a solution may need to be deployed prior to an Integrated solution being available, and that the EFB_Aft condition is least appropriate, the question is whether a clear choice can be made between the EFB_Aft+AGD and the EFB_Fore conditions. In addition to those measures that were insensitive to any changes in Avionics Condition, there were other measures that failed to show any significant differences in pairwise comparisons between these two remaining retrofit solutions. Differences between the EFB_Fore and EFB_Aft+AGD were not statistically significant for: post-run ratings regarding the degree to which they are operationally supportive of IM, number of speed deviations, awareness of speed profile deviations, crew noticing time, and heads-up time.

Of these two, more of the remaining assessments favored the EFB_Fore condition. When using the EFB_Fore condition, crews sampled the FIM-related display more than when using the EFB_Aft+AGD condition (though both were less than the Integrated condition, and both were better than the EFB_Aft condition). Although not statistically different in terms of noticing time, crews with the EFB_Fore condition entered speeds in the MCP more quickly and required fewer reminders. In addition, pilots' ratings with respect to awareness of new speed commands were more complimentary when using the EFB_Fore condition. The converging evidence that these results provide is buttressed by pilots' post-experiment assessment of operational acceptability for supporting these IM operations. Whereas ratings for the EFB_Fore condition were significantly higher (indicating more supportive) than the midpoint of the rating scale, ratings for the EFB_Aft+AGD condition could not be shown to significantly differ from the midpoint ("borderline" acceptable). However, the inclusion of the AGD did seem to improve visual access to FIM information, as the EFB_Aft+AGD condition was rated less distracting than the EFB_Fore condition and dwell times were significantly shorter. Finally, whereas the EFB_Fore condition was associated with fewer speed profile deviations, the EFB_Aft_AGD was associated with less extreme speed deviations. One could surmise that this finding derives from the improved performance in detecting and entering speed commands in the first place, when using the EFB_Fore condition; but that if a deviation is left unattended to, the indications provided in the AGD better supported management back to the speed profile.

Pilots were asked to compare pairs of Avionics Conditions on the degree to which they supported IM operations. As reported above, the Integrated condition was most preferred, and the EFB_Aft least preferred. Of the remaining retrofit solutions, then, which is preferred?

Closer inspection of results show that the Notification Method pilots experienced affected their ranking of the remaining two Avionics Conditions. When crews experienced aural indications for all IM events (i.e., the AAA method), they were more likely to prefer the EFB_Fore condition than the EFB_Aft+AGD condition. When crews had the VAV or VVV method, their second choice was the EFB_Aft+AGD condition. The following section reviews main effects of Notification Methods and other interactions of these with Avionics Condition.

6.3 Notification Methods

Comparative comments were not available regarding Notification Methods, as each crew experience only one throughout all their runs. This section is therefore based on comparisons of the groups of pilots that experienced each of the three Notification Methods.

A subset of 10 pilots (four from the VAV method, and six from the AAA method) were asked to rate the degree to which the specific tone used in this study was annoying, and to rate the perceived urgency it conveyed. Pilots felt the aural indication was appropriate: it received central ratings with regard to perceived urgency, and below central responses when rated on annoyance. It appeared to be salient enough to capture attention in most cases, but not so annoying that it was impossible to conduct other flight related tasks. Because these data are imbalanced and not unbiased with respect to Notification Method, these results are only discussed here as anecdotal evidence that the aural indication was acceptable and appropriate in terms of its perceptual qualities.

As expected, most of the scenario assessment questions (i.e., the acceptability of speedbrake use and speeds, the frequency of commanded speeds, and the degree to which these were consistent with expectations) were impacted by other events, and didn't conflict with other information in the scenario. However, Notification Method manipulations showed no influence on oculometer measures (FIM sampling, heads-up sampling, noticing time of new commanded speeds, dwell times on the first FIM-related look after a new speed, the NNI workload measure, or the KLD scan path comparison measure). Ratings of peak and average workload as assessed by the MCH scale were unaffected, as were ratings for the NASA-TLX scales for perceived effort and performance, temporal and mental demand. SART ratings of situation awareness also did not differ over these Notification Methods, and vertical path deviations were similar.

Of the remaining evidence, the primary finding was positive with respect to the AAA method, and less so for the VAV and VVV methods. The AAA method was found superior to both the VVV and VAV methods (for all of the Avionics Conditions) in terms of: the number of reminders necessary, subjective assessment of ability to detect speed deviations, and NASA-TLX composite workload scores. For the EFB_Aft condition, the AAA method received better ratings than either the VVV or VAV methods in terms of lower distraction, and better ratings of situation awareness with respect to new speed commands as well as other aspects of flight. For both the EFB_Aft and Integrated conditions, the AAA method resulted in fewer speed profile deviations than the other two methods. For some measures, pilots with the VAV method behaved similarly to those with the AAA method and both these methods were more supportive than the VVV method. Principally, this was evidenced in the time to enter commanded speeds in the MCP – and this was true across all Avionics Conditions. For the Integrated and EFB_Aft conditions, using either the AAA or VAV method resulted in fewer speed profile deviations; and for the EFB_Aft condition, also less extreme speed profile deviations.

For the other two conditions (EFB_Fore and EFB_Aft+AGD), the AAA method was clearly more supportive of adhering to the speed profile than the VVV method. Other measures demonstrated this strong benefit of the AAA method compared to the VVV method, with the VAV method indistinguishable from either. The AAA method was more supportive than the VVV method, if not the VAV method, as evidenced by its main effect over all Avionics Conditions in terms of: fewer conformance deviations; faster times to enter speeds in the MCP; less frustrating; and better self-reported situation awareness of new commanded speeds, the speed profile, as well as non-FIM aspects of flights. In addition, for both the Integrated and EFB_Aft conditions, the AAA method was rated significantly less distracting than the VVV method (perhaps taken as less necessary to actively monitor the FIM display); and for the Integrated condition, the AAA method supported significantly faster noticing of new speeds than did the VVV method.

The utility of and preference for the AAA method were reinforced by commentary and post-experiment questions that asked crews to consider whether and which IM events should have aural annunciations. Twenty-two of the 24 pilots indicated that an aural annunciation would be helpful for at least one of the IM events; and half of the pilots indicated that they should be used for all three IM events. Note that this includes pilots who did not receive aural indications during

88

the test. All eight of the pilots in the VVV method expressed an interest in having an aural indication for at least one IM event. When asked to provide ratings for the appropriateness of aurals per IM event, pilots who experienced the AAA method were most enthusiastic about using aural indications, and had high ratings for all three IM events. The pilots who experienced the VAV method suggested that they would appreciate an aural indication as a reminder, and pilots who experienced the VVV method desired an aural indication for command speed onset. The finding that no method group's ratings were highest for using aural indications for the conformance deviation indicator coincides with the fact that no pilots suggested it in open format responses.

Interestingly, pilots who experienced aural indications in the context of the AAA method gave high ratings to the appropriateness of aural indications, whereas those pilots who experienced aural indications which alerted conformance deviations (the VAV method) rated them lowest. This difference could be hypothesized to be due to not the aural indication itself, but what it means to receive one, and what might be an affective response to such an occurrence. In the case of the AAA method, the most common occurrence of an aural indication is notification that a commanded speed has occurred. In the case of the VAV method, the only use of the aural indication is an alert that the pilot has clearly erred. An alternative hypothesis for this finding is that the AAA method habituates the crew to the presence of these aural notifications, and in an otherwise fairly sterile flight deck (perhaps these pilots were less chatty in an experimental setting than they might be in real operations), the fact that there are more of these for these indications maintains engagement in the IM operation; whereas the aural indications in the VAV method the crew has clearly lost SA. This hypothesis garners some support from the fact that the AAA method was actually rated as less distracting than the other methods.

To contextualize the findings with respect to the use of aural indications for IM events, it is appropriate to review the conditions under which these occurred. Notifications occurred for three IM events: commanded speed onset, reminder to implement a commanded speed, and to indicate a conformance deviation. Pilots' suggestions for reminder thresholds were consistent with the manner in which they were implemented in this study (after 10 seconds of non-response). Pilot recommendations for conformance deviations were also fairly consistent. For these, there was a broad spread of suggested time and speed thresholds. The pilots were more likely to consider the indications for conformance deviations (a two second delay, then occurring

when over 7 knots off for over 12 seconds and not converging to the speed profile) to be false alarms than the reminders (this finding was most obvious for pilots who used the VAV method).

6.4   Pilot Flying Role and Location

The primary focus of this study was to address significant impacts of Avionics Conditions and Notification Methods. Each pilot experienced both PF and PM roles, as a means of providing variety and hopefully engagement in the study. When switching roles, pilot participants remained in same seat. Therefore, for measures that pertain to the crew (rather than each pilot), the effective manipulation was Seat (location), rather than Role. Creating this manipulation necessitated its consideration in analyses, and so these were included as exploratory. Most ratings associated with scenario design (realism, comfort with speeds) were not affected by pilot role (Pilot Flying (PF) or Pilot Monitoring (PM)), nor were most measures aimed at addressing differences in Avionics Configurations. Measures that showed no significant effects of Role (or Seat), nor interactions of these factors with Avionics Conditions or Notification Methods include: speed excursion magnitude and number; number of reminders; SART attentional demand subscale and total score; ratings associated with new commanded speed, speed profile deviation and TTF awareness; degree of distraction, frustration, and workload experienced in scenarios; FIM display and Head-up sampling; Dwell Times; and ratings of avionics acceptability for supporting IM operations.

The Pilot Flying, for a given sector of the flight, is directly responsible for flying the aircraft. The Pilot Monitoring monitors the PF and carries out support duties such as communications and check-list reading. With the increased responsibility of the flight, it is not surprising that when pilots acted as PF, they reported higher workload ratings in terms of NASA-TLX subscales for mental demand, temporal demand, effort, frustration, and the composite score; as well as in MCH average and peak workload ratings. As PFs are directly responsible for flying the aircraft, they are the most direct consumers of new commanded speeds. PFs were significantly faster to notice new commanded speeds, but only by approximately 200msec. Investigation of noticing times for each pilot revealed a significant interaction with Notification Method. When pilots received only visual indications of new speed commands, PMs were statistically more likely to be the first to notice these than when an aural accompanied a new commanded speed.

Ratings on the SART subscale associated with "understanding" were higher when pilots were PF than PM, but scores pertaining to pilots' ability to maintain situation awareness of non-IM aspects of the flight were higher when PM than PF. Together, these results seem inconsistent, but the subtle difference may hinge on the first relating to the quality of the situation awareness, whereas the latter referred to the excess capacity (perhaps monitoring ability) to acquire situation awareness, and the broader responsibility PMs have for monitoring. When serving in the role of PM, pilots provided higher ratings for the acceptability of the commanded speeds, the frequency of the speeds as they occurred in scenarios, and the degree to which these were consistent with expectations – consistent with a lower assessment of workload in general. The oculometer NNI metric is higher for PM than PF, indicating a more systematic scan pattern; theoretically induced by greater workload. This last finding is both inconsistent with results from subjective ratings of workload and counter-intuitive with respect to PF and PM duties.

While ratings could be assessed in terms of pilot role, because both the PF and PM could provide data from those perspectives, performance measures taken at the crew level were assessed in terms of whether the PF was flying from the left seat or the right seat. There were no significant main effects for this factor, but in some cases, Avionics Condition or Notification Method interacted with which seat had the PF. For the AAA method, entering speeds in the MCP was fastest when the PF was in the left seat; but when using the VAV method, faster when the PF was in the right seat. When using the AAA method, vertical path deviations were more extreme when the PF was in the left seat; and when using the Integrated condition, more extreme when the PF was in the right seat.

When the PF was in the left seat (as is typical) and used the AAA method, there were more extreme vertical excursions, but the crew entered speeds in the MCP more rapidly than when the PF was in the right seat. In contrast, when using the VAV method, speeds were entered more quickly if the PF was in the left seat. The only interaction with Avionics Condition implies that if using the Integrated condition, there are more extreme vertical excursions when the PF is in the right seat. Crews in this study were pilots from the same airline, and for most crews, contained one captain and one first officer. Typically, the captain would sit in the left seat, and the first officer in the right seat. As such, the findings associated with the Seat may include influences not only of location, but also experience.

7 Discussion

Results of this study can be used for several purposes: (1) as design guidance for avionics that effectively support Flightdeck Interval Management; (2) as guidance for, and to identify issues pertinent to certification of, and flight standards for FIM operations; and (3) to offer insights to those who will conduct similar evaluations in the future. Observations from this study are considered for each purpose.

7.1 Designing & Evaluating FIM-Supportive Avionics

First, with the exception of the EFB_Aft condition, all conditions were at least somewhat acceptable in terms of rating data, and deviations on measures were not extreme. The results agree with prior research, where situation awareness and workload ratings are acceptable (Bone, 2008; Swieringa, 2011). However, if a choice is to be made on the case of effectiveness, the preponderance of evidence seems to suggest the following rank order of the Avionics Conditions tested (best to worst): (1) Integrated, (2) EFB_Fore, (3) EFB_Aft+AGD, and as a distant fourth, the (4) EFB_Aft condition. One should pause to consider whether all evidence ought to be equally considered. While the vast majority of ratings and rankings seem to consistently favor the Integrated condition, at least one objective measure – the number of out-of-conformance indications – tells a different story. For this measure, the Integrated condition resulted in the worst performance, particularly when no aural indications accompanied it. In the Integrated condition, speed profile deviation information was presented graphically as a difference between current indicated airspeed and a symbol on the airspeed tape. This information was also provided as a fast/slow numerical value on the MCDU IM page, however as pilots often selected another MCDU page during the approach, the information in this form was not always visible. As such, the finding that the Integrated condition results in more speed profile deviations, particularly when no aural indications are provided (and when pilots are clearly sampling the FIM-related display and noticing and entering commanded speeds) suggests that the indicator for speed profile deviation is not salient enough to support its use in this display – and pilots said as much. Integrating FIM information into the PFD would seem beneficial in that this minimizes the need to sample an additional display for FIM information, and consolidates FIM speed information with current speed information. However, it is challenging to detect changes in cluttered displays (Moacdieh & Sarter, 2014), and may especially be difficult when pilots would be attending closely to other elements in this display as

a result of well-developed scan patterns. It is an open question as to whether additional exposure, or a revised PFD design, would improve these results and put them in accord with subjective impressions and ratings. Intuitive display concepts are desirable in order to minimize training, both in terms of costs and retention of the training; so if IM is conducted infrequently, it can be still be easily used. Increasing salience of this element would come at a cost to noticing other facets of the display, and redesign of this display in particular should be supported by significant research.

It isn't surprising that the EFB_Aft condition was rated poorly due to its aft location in the flight deck, but it is notable that when aural indications accompanied it, some measures of effectiveness improved – ostensibly because these annunciations cued pilots to visually access the display, and thereby reduced monitoring requirements and the probability of missing speed deviations and possibly new commanded speeds. The AGD was designed to solve this same problem, minimizing the need to sample the EFB by repeating essential FIM information in the forward field of view. Before being exposed to the concept, pilots felt the AGD was a good solution; however, after experiencing the condition, the AGD was not as well-received. Pilots' ratings of operational acceptability for the EFB_Aft condition were low, but the addition of the AGD only resulted in a "borderline" average rating). Several pilots mentioned that they thought that they would prefer the condition with the AGD, but found that its indication of new commanded speeds lacked salience, and so reliance on it resulted in missing target speeds. Performance metrics were mixed. The EFB_Aft+AGD was associated with less extreme speed excursions than the EFB_Aft condition, but failed to improve response to new speeds (i.e., the time to enter the speeds in the MCP and number of reminders). As currently designed, this configuration appears to not be sufficiently supportive. One might suspect that this is in part due to the different accommodation planes for this instrument and the main foci of attention – the instrument panel and the out-the-window view. While the MCP is at the same focal length, this instrument is not generally fixated much during the arrival phase (Mumaw, Sarter & Wickens, 2001), and when it is, this behavior is likely endogenously driven by task requirements.

The EFB_Fore condition seems to be appropriately supportive, but placement is key to its successful implementation. Placement should ensure a comfortable posture for manual entry, and should ensure that (in combination with display formatting) presented items are easily viewed and legible at a comfortable accommodation distance. In the implementation for this

experiment, EFB display content was the same for the left and right seat. However, this experiment revealed this was not optimal. When indications were presented in the upper right corner of the display, they were noticed by left seat pilots (as they were within peripheral vision), but could more easily be missed by right seat pilots. Future research should pursue the characterization of salience with respect to not only viewing angle of the display device, but also the degree to which the display features exogenously direct attention.

While this study provides no data to support this conclusion directly, one might infer that the best retrofit solution might have an EFB_Fore condition augmented with profile deviation indications in the forward field of view. However, if tested, the EFB_Fore+AGD condition should benefit from redesign in accord with pilots' concerns regarding ergonomic placement of the EFB, and salient indications on a forward-field-of-view display.   Other concepts that were not assessed in this study should also be considered – to include presentation of FIM information on heads-up displays.

The preference for aural indications by these crews and their recommendations to have a separate aural indication for each type of IM event runs counter to concerns regarding proliferation of aural indications on modern commercial flight decks (Boucek, Hanson, Berson, Leffler, Po-Chedley & Hendrickson, 1981). While subjective data has been known to dissociate from performance data (Andre & Wickens, 1995); here, use of aurals was broadly associated with both higher ratings and performance enhancement. Further, the AAA method (with aural indications for new commanded speeds, conformance deviations and reminders) often showed significant benefit over the VAV method (where an aural indication is provided only for speed profile conformance deviations). Pilot comments may explain why - the clear preference was for an aural indication for new commanded speeds. Apart from the 12 pilots who desired aurals for all IM events, 10 pilots wanted an aural indication (alone or in combination) for commanded speeds versus only five for reminders and four for conformance deviations. A potential source of confusion existed in the AAA condition; however, if a crew experienced a conformance deviation, was in the process of correcting back to the profile, and received another aural indication, they occasionally mistook this as another notice of the conformance deviation, when it actually signaled a new commanded speed. Future research should investigate the tradeoff between single versus multiple tones for the three different IM events.

The deleterious effects of false alerts in terms of reducing compliance, are well-known (Geels-Blair, Rice, & Schwark, 2013), and should motivate more sophisticated consideration of the logic that triggers reminders and speed deviation indications. In particular, different speed deviation thresholds may be appropriate for different regions of flight, as would be consistent with pilots and system expectations.

7.2   Aircraft Certification and Flight Standards Considerations

Avionics Characteristics

With respect to certification of FIM avionics, this study finds evidence that while the Integrated solution tested here seems to best support flying FIM-commanded speeds, the tested EFB_Fore condition was acceptable. Both conditions require additional design consideration. In the case of the Integrated solution, redesign should ensure that speed deviation information is salient, but not intrusive to other uses of the PFD, and that vertical path deviation information is integrated into the MCDU IM page to permit simultaneous monitoring of speed deviation and path deviation. In the case of the EFB_Fore condition, redesign should ensure that its placement ergonomically supports comfortable viewing and data entry, and consider augmentation with a redesigned forward-field-of-view display that facilitates detection of speed commands, and especially speed profile excursions. Other display concepts (e.g., heads-up display of FIM information) are possible, and future research should consider these as well.

In this study, the EFB provided FIM information for retrofit solutions; and for all solutions also contained approach charts and the airport diagram. Paper charts were also provided. The experimental team reasoned that EFBs would contain charts before FIM operations were implemented and that the management of different content on a single auxiliary device could present a concern (c.f. Chandra, Yeh, Riley & Mangold, 2003). In implementing this EFB design, we noted the need to minimize actions to navigate between charts and the FIM display, and to ensure that when new FIM information was available (a new speed, reminder or conformance indication), that this was visible to pilots even when not on the FIM page. We implemented a solution in which the top right corner included a message that FIM information was available, if pilots were on a chart page. An exploratory case was conducted in which this indication was augmented with the actual information. Some pilots commented favorably on this, but most didn't notice it – largely because crews had developed a strategy by which the

FIM display was available on one EFB, and charts on another. For pilots in the right seat, simply providing the same display format put it further in the periphery (by the width of the EFB) than it was for pilots in the left seat. This is an interesting implication for display design; mirrored formats would ensure that most important information is inboard, but this would negatively impact use for those who are exposed to this from both left and right seats. Further research is required to study conduct of FIM operations with an auxiliary display that is as fully functional as it will be for pilots in actual operations.

Aural indications received strong support in this study. However, scenarios did not include many aural distractors that could occur in actual flight operations. In addition, because Notification Method was a between-subject variable, this study provides no direct comparisons within the same individual across the tested methods. Further investigation is appropriate to determine not only the appropriate conditions under which to cue FIM events with aural annunciations, but to design the appropriate aural indication for this purpose. While not studied here, some subjects recommended the use of voiced annunciations to support FIM operations, and this too requires additional research. This research found aural indications to significantly benefit all the retrofit solutions in a variety of ways. However, the EFBs tested (and implemented in many aircraft) currently do not have the capability to aurally annunciate. Should subsequent research confirm the advantages of aural indications to the point that these are widely accepted as necessary, this would necessitate consideration of EFB design, design of a confederated annunciation system, and exploration of how auxiliary devices with aural indications would be managed with other flightdeck aural indications.

Assessing FIM Avionics

The dependent measures taken in this study suggest some appropriate considerations for scoring a candidate FIM avionics solution in terms of supporting the task of flying FIM speeds. While not studied here, FIM avionics must also support clearance entry and decisions with regard to off nominal situations. In this broader consideration, when assessing an avionics solution for FIM, one should consider the degree to which it supports IM operations by:
- supporting flight path management and comfortable speed management;
- permitting spare workload capacity to address other aspects of the mission;
- minimizing additional workload;

- permitting integration of FIM information into an efficient and familiar scan pattern;
- supporting efficient and error-proof data entry of FIM clearance information;
- providing effectively alerts to exogenously redirect attention as necessary;
- supporting timely detection of new commanded speeds;
- supporting extraction of information quickly and unambiguously;
- supporting timely detection of speed profile conformance deviations;
- providing sufficient notice of profile excursions to avoid significant deviation and such that these are not considered false alarms;
- supporting effective and timely decisions for IM cancellation and resumption;
- providing well-timed reminders to facilitate speed command entry in a timely manner;
- being consistent with the flightdeck design philosophy, cockpit resource management (CRM) and communication conventions;
- not being frustrating, distracting, or awkward to use.

7.3   Methodological Considerations & Caveats

Measures

It is good practice to ask some questions one hopes are not sensitive to experimental manipulations. These include checks on the sufficiency of training, and the realism of the simulation environment and scenario elements and pacing. Assessment of the concept of IM operations was largely addressed in post-experiment open format questioning, but also specifically addressed by questions about the acceptability new elements introduced by this operation (flying commanded speeds, and the tasks of attending to these and speed profile adherence). Finally, it is important to assess the effects of what one is testing on all the rest (obviously it's of little use to optimize FIM performance at the expense of all else) – in this study, this was obtained by asking about SA for other aspects of the flight, and assessing Heads-Up time.

This study assessed constructs hypothesized to be affected by Avionics Configurations through performance data, post-run questionnaire instruments and Likert items, and post-experiment pairwise comparisons, Likert and forced-choice items, and open ended questions. For some constructs, we chose to ask several similar items. For the purposes of the interpretation of participant's experiences, this approach hopes to demonstrate converging evidence. In

addition, as this study was a precursor to a flight demonstration, an objective was to help refine an efficient set of question for this more constrained environment in order to maximize sensitivity to pilots' experiences. Indicators that were sensitive to both Avionics Condition and Notification Method indicators included performance measures, post-run questionnaire items, and post-experiment questions. In terms of objective measures, Avionics Configurations were best distinguished by: the number of reminders and conformance deviations, and the time to notice and enter the commanded speeds in the MCP. In addition, manipulation of Avionics Condition (but not Notification Method) was sensitive to the magnitude of speed excursions in the period following display of new speed commands.

Ratings on post-run questionnaire items that were most affected by differences in Avionics Configurations included: three questions addressing detection of commanded speeds ("*I was never surprised to notice that the speed target had changed*," "*I was aware of commanded speed changes in an appropriate timeframe*," "*Avionics and annunciations appropriately supported my ability to detect speed targets in a timely manner*"); one question regarding speed profile deviation awareness ("*Avionics and annunciations in this scenario supported timely deviation detection*"); a question regarding the degree of speedbrake required ("*The speeds given did not require an uncomfortable level of speedbrake use*"), and a general question regarding acceptability ("Use of the avionics and annunciations are acceptable for conducting IM operations").

Commonly used instruments were less sensitive than those above. SART showed differences between Avionics Conditions, but not Notification Methods; the NASA-TLX composite score and the frustration subscale were sensitive to Notification Methods, but not Avionics Condition; and the commonly-used MCH workload scale was insensitive to both Avionics Conditions and Notification Methods – both when assessing average and peak workload of a scenario. Oculometer measures (FIM Sampling, Heads-Up sampling, Pilot Noticing Time, Minimum Crew Noticing Times, and Dwell Times) were primarily sensitive to manipulations in Avionics Conditions – and, only for Noticing Times, also for the interaction of Avionics Condition and Notification Method. Neither the oculometer workload measure that is based on scan regularity (NNI), nor the KLD measure of PF and PM scan coherence showed differences. It's important to note that oculometer data provides POG – point of gaze; but looking is not always seeing. As such, it is important to employ sensitive selection rules based on data quality, number of missed frames, and windowing on important events to ensure useful results. The saving grace of

oculometery is that, despite these measurement errors, findings gain credence from the sheer volume of data provided.

While not a focus of this study, comparisons of PF and PM behavior and ratings showed some significant results, including significant interactions with Avionics Conditions and Notification Methods. Future research that aims to address CRM impacts of FIM may find it useful to consider these measures in particular: magnitude of vertical error, time to notice commanded speed changes and time to enter enter speed in the MCP. Interestingly, standard instruments were more sensitive to differences in Role, than experimental conditions. PF and PM ratings significantly differed on SART, NASA-TLX, and MCH (peak and average workload) ratings. The NASA-TLX and MCH instruments showed differences in workload levels, and these were also reflected in questions regarding the acceptability of scenario difficulty (speedbrake required, frequency and acceptability of speeds, and ability to support SA of other aspects of the flight).

Study Generalizability and Caveats

Pilots were generally complimentary of the IM concept and were able to conduct these operations fairly well with minimal training. To ensure that the experiment did not introduce interfering artifacts, a set of post-run questions specifically addressed the acceptability of scenarios (e.g., commanded speeds, and the frequency of commanded speed changes as presented were: operationally acceptable, safe and appropriate, didn't conflict with other information, their timing was consistent with expectations, and were comfortable to fly; the level of speedbrakes required was not uncomfortable, and the events in the scenario were realistic). These received generally high ratings, and were mostly insensitive to Avionics Condition and Notification Method. Post-experiment ratings associated with training effectiveness and the realism of the simulation environment and the scenarios were all positive. In sum, these findings support generalization of these results to actual operations.

However, some important caveats are required. The reader is to be reminded that the results obtained here are inextricably linked to the particular implementations under test. For example, it was found that certain visual indications lacked salience, and results may differ if the design of the visual indications were altered. As with most simulation studies, sample size is a concern. In this study, it is important to note that the Notification Method manipulation is a between-

subject variable, and so received fewer data points per subject than the within-subject variable testing Avionics Conditions – and no subject was able to directly compare these methods. This study addressed the effectiveness of different Avionics Configurations in supporting FIM operations during execution phase – and so assessed acceptability of only a subset of the IM concept of operations. Consideration of task requirements associated with initiation and terminating/suspending FIM are necessary to fully appreciate the degree to which an avionics configuration supports the full IM concept of operations. Finally, this study presents data for participants that were newly exposed to FIM operations, and self-selected to participate. A longitudinal investigation with a broader set of participants, in actual operations will certainly show a both a wider variety of, and the time course for tuning, FIM-related performance.

7.4   Future Research

The findings from this study strongly suggest the benefits of the Integrated Avionics Configuration and the acceptability of at least one retrofit avionics solution. However, in both cases, additional research is required to optimize display design in both integrated and retrofit approaches to ensure supportive salience of required display elements. This study also found compelling advantages when FIM avionics were augmented with aural indications. However, additional research is required to: (1) confirm such findings in a study that varies Notification Method within subjects – to ensure that findings here were not an artifact of dissimilar groups; (2) confirm findings in a richer auditory context (i.e., with off-nominal indications); (3) determine the best allocation of aural indications to FIM events (based on commentary, this study may not have tested some viable alternate codings) and whether all events should have a single, or differentiated annunciations; (4) determine the best formatting characteristics for an aural indication (tonal characteristics, or voiced).

Pilots generally commented that FIM didn't require specific CRM training, and fit well with current procedures. However, we did observe that crew members evolved different strategies for supporting each other in monitoring for new commanded speeds and ensuring adherence to the commanded speed profile. This, with the finding that Role and sometimes Seat (which was somewhat confounded with experience) significantly interacted with Avionics Conditions and Notification Methods, leads us to suggest that CRM factors in FIM operations should also be the subject of future study.

8   References

Abbott, T. S.; and Moen, G. C, (1981). Effect of Display Size on Utilization of Traffic Situation Display for Self-Spacing Task, *NASA TP-1885*. NASA Langley Research Center: Hampton, VA.

Abbott, T.S. (2012). *An Overview of a Trajectory-Based Solution for En Route and Terminal Area Self-Spacing: Third Revision*. NASA/CR-2012-217786. NASA Langley Research Center: Hampton, VA.

Andre, A.D. & Wickens, C.D. (1995). When users want what's not best for them. *Ergonomics in Design*, *October*, 10-14.

Bailey, R., Kramer, L., & Prinzel, L. (2006). Crew and display concepts evaluation for synthetic/enhanced vision systems. In *Proceedings of the SPIE, Enhanced and Synthetic Vision Conference.*

Baxley, B. T., Swenson, H. N., Prevot, T., & Callantine, T. J., (2012). NASA's ATM Technology Demonstration-1: Integrated Concept of Arrival Operations, *31st Digital Avionics Systems Conference (DASC),* Williamsburg, VA, October 2012.

Baxley, B.T., Swenson, H.N., Robinson, J.E., Prevot, T., Callantine, T.J., Scardina, J., & Greene, M. (2012). Air Traffic Management Technology Demonstration-1 Concept of Operations (ATD-1 ConOps). *NASA/TM–2012-217585*. NASA Langley Research Center: Hampton, VA.

Becker, S.I. (2011). Determinants of Dwell Time in Visual Search: Similarity or Perceptual Difficulty?   *PLoS ONE 6*(3): e17740. doi:10.1371/ journal.pone.0017740

Bone, R. S., Helleberg, J., Domino, D. A., & Johnson, N. (2003). Effects of Traffic Display Size and Location During an Instrument Approach Spacing Task: Cockpit Display of Traffic Information (CDTI) Enhanced Flight Rules (CEFR) Simulation 2. *MTR 03W0000030,* The MITRE Corporation Center for Advanced Aviation System Development: McLean, VA.

Bone, R. S., Penhallegon, W.J., & Stassen, H.P. (2008a). Flight Deck-Based Merging and Spacing (FDM) Impact on Flight Deck En Route Operations: FDMS 2 Simulation. *MTR080134.* The MITRE Corporation Center for Advanced Aviation System Development: McLean, VA.

Bone, R. S., Penhallegon, W.J., & Stassen, H.P. (2008b). Flight Deck-Based Merging and Spacing during Continuous Descent Arrivals and Approach: Impact on Pilots - FDMS 3 Simulation. *MTR080208.* The MITRE Corporation Center for Advanced Aviation System Development: McLean, VA.

Boucek, G. P., Hanson, D.C. Berson, B.L., Leffler, M.F., Po-Chedley, D.A. & Hendrickson, J.F. (1981). Aircraft alerting systems standardization study, *AIAA/IEEE Digital Avionics Systems Conference*, 119-128. AIAA: New York, NY.

Callan, D. J. (1998). Eye movement relationships to excessive performance error in aviation. In *Proceedings of Human Factors and Ergonomics Society Annual Meeting*: Santa Monica, CA.

Callantine, T. J., Cabrall, C. D., Kupfer, M., Omar, F. G., & Prevot, T. (2012). Initial Investigations of Controller Tools and Procedures for Schedule-Based Arrival Operations with Mixed Flight-Deck Interval Management Equipage, *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Paper 2012-5673. AIAA, Indianapolis, IN.

Camilli, M., Terenzi, M., & DiNocera, F. (2007). Concurrent validity of an ocular measure of mental workload. In D. deWaard, G.R.J. Hockey, P.Nickel, and K.A. Brookhuis (Eds.), *Human Factors Issues in Complex System Perform*ance. Shaker Publishing: Maastricht, the Netherlands.

Casali, J.G. & Wierwille, W.W. (1983). A comparison of rating scale, secondary task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors, 25,* 623-642.

Chandra, D., Yeh, M., Riley, V. & Mangold, S. (2003). Human factors considerations in the design and evaluation of electronic flight bags (EFBs) – version 2.   *DOT/FAA/AR-03/67, DOT-VNTSC-FAA-03-07.* Office of Aviation Research: Washington, DC.

Comstock, J., Harris, R., Coates, G. & Kirby, R. (1987). Time-locked time-histories - A new way of examining eye-movement data. International Symposium on Aviation Psychology, 4th, Columbus, OH. Columbus, OH.

Cooper, G. E & Harper, R. P. (1969). The use of pilot rating in the evaluation of aircraft handling qualities. *ASD-TR-76-19*. NASA Ames Research Center: Moffett Field, CA.

Credeur, L. (1977). Basic analysis of terminal operation benefits resulting from reduced vortex separation minima. *NASA/TM-78624.* NASA Langley Research Center: Hampton, VA

Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R. & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavioral Research Methods 44*(4):1079-100.

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

Di Nocera, F. (2007). Cognitive Aspects and Behavioral Effects of Transitions Between Levels of Automation. *European Office of Aerospace Research & Development Report* (Contract Number FA8655-05-1-3021).

Ellis, K.K.E., Arthur III, J.J., Latorella, K. A., Kramer, L., Shelton, K J., Norman, R.M. & Prinzel, L.J. (2012). Quantifying Pilot Visual Attention in Low Visibility Terminal Operations. In *Proceedings of the AHFE International Conference on Applied Human Factors and Ergonomics*: San Francisco, CA.

Federal Aviation Administration (2013). *FAA Aerospace Forecast, Fiscal Years, 2013-2033* http://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=14374 (accessed 8/3/2013)

Federal Aviation Administration (2013). *NextGen Implementation Plan June 2013*, https://www.faa.gov/nextgen/implementation/ (accessed 8/23/13).

Federal Aviation Administration (2010). *Advisory Circular: Flight Crew Alerting* (AC 25.1322-1)

Geels-Blair, K., Rice, S., & Schwark, J. (2013). Using System-Wide Trust Theory to Reveal the Contagion Effects of Automation False Alarms and Misses on Compliance and Reliance in a Simulated Aviation Task. *The International Journal of Aviation Psychology. 23*(3), 245-266.

Gonzalez, C., Lewis, B.A., Roberts, D.M., Pratt, S.M., & Baldwin, C.L. (2012). Perceived urgency and annoyance of auditory alerts in a driving context. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*: San Francisco, CA.

Harden, J.W. & Latorella, K.A. (2014). *Software for the analysis of Smarteye Oculometer Data for Workload and Scan Path Comparions.* Manuscript in preparation.

Harris, Sr., R.L., Glover, B.J. & Spady, Jr. A. (1986). *Analytical Techniques of Pilot Scanning Behavior and Their Application.* NASA-TP-2525 19860018448. NASA Langley Research Center: Hampton, VA.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) Human Mental Workload. North Holland Press: Amsterdam.

Hébraud, C., Hoffman, E., Pene, N., Rognin, L., & Zeghal, K. (2004). Assessing the Impact of a New Air Traffic Control Instruction on Flight Crew Activity, *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Paper 2004-5104. AIAA: Providence, RI.

Kullback, S. & Leibler, R.L. (1951). On information sufficiency. *Annals of Mathematical Statistics 22*, 79-86.

Latorella, K., Ellis, K., Lynn, W., Frasca, D., Burdette, D., Feigh, C. & Douglas, A. (2010). Dual Oculometer System for Aviation Crew Assessment. *54th Annual Meeting of the Human Factors and Ergonomics Society*: San Francisco, CA.

Leahy, J. (2012*). Navigating the Future: Global Market Forecast 2012-2031*. Airbus. http://www.airbus.com/company/market/forecast/ (accessed 8/3/2013)

Levitt, I. & Weitz, L.A. (2011). Towards Defining Required Interval Management Performance. *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM)*

Lohr, G. W., Oseguera-Lohr, R. M., Abbott, T. S., & Capron, W. R. (2003**).** Flight evaluation of a time-based airborne inter-arrival spacing tool. *5th USA/Europe ATM R&D Seminar*. Budapest, Hungary.

Moertl, P.M. & Pollack, M.E. (2011). Airline based en route sequencing and spacing field test results: Observations and lessons learned for interval management. *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011):* Berlin, Germany.

Mumaw, R.J., Sarter, N.B., & Wickens, C.D. (2001). Analysis of pilots' monitoring and performance on an automated flight deck. *International Symposium on Aviation Psychology*: Dayton, OH.

Murdoch, J. L., Wilson, S. R., Hubbs, C. E., & Smail, J. W. (2013). Acceptability of Flight deck-based Interval Management Crew Procedures, *AIAA Modeling and Simulation Technologies Conference and Exhibit*. AIAA: Boston, MA.

Oseguera-Lohr, R.M., Lohr, G.W., Abbott, T.S., & Eischeid, T.M. (2002). Evaluation of Operational Procedures for Using a Time-based Airborne Inter-arrival Spacing Tool. *Proceedings of the AIAA Aircraft Technology, Integration, and Operations Conference*, Paper 2002-5824. AIAA: Los Angeles, CA.

Oseguera-Lohr, R.M., Lohr, G.W., Abbott, T.S., Nadler, E.D. & Eischeid, T.M. (2005). Evaluation of a Tool for Airborne-Managed In-Trail Approach Spacing. *NASA/TP-2005-213773*. NASA Langley Research Center: Hampton, VA

Penhallegon, W.J., Mendolia, A.S., Bone, R.S., Orrell, G.L. & Stassen, H.P. (2011). Flight Deck-Based Interval Management-Spacing During Departures: Flight Crew Human-In-The-Loop Simulation. *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011):* Berlin, Germany.

Prevot, T., Callantine, T., Kopardekar, P., Smith, N., Palmer, E., & Battiste, V. (2004). Trajectory-oriented operations with limited delegation: an evolutionary path to NAS modernization. *AIAA 4th Aviation Technology, Integration and Operations (ATIO) Forum, Paper 2004-6449.* AIAA: Chicago, IL.

Prevot, T., Callantine, T. J., Homola, J. R., Lee, P. U., Mercer, J. S., Palmer, E. A. & Smith, N. M. (2007). Effects of Automated Arrival Management, Airborne Spacing, Controller Tools, and Data Link,, AIAA-2007-6554*, AIAA Guidance, Navigation and Control Conference and Exhibit,* Hilton Head, SC.

Prevot, T., Baxley, B. T., Callantine, T. J., Johnson, W. C., Quon, L. K., Robinson, J. E., & Swenson, H. N. (2012). NASA's ATM Technology Demonstration-1: Transitioning Fuel Efficient, High Throughput Arrival Operations from Simulation to Reality, *International Conference on Human-Computer Interaction in Aerospace (HCI-Aero):* Brussels, Belgium.

Redeborn, B. (2013). Next Generation Airspace Developments: Europe's Contribution to Global Interoperability. *Integrated CNS (Communication, Navigation & Surveillance) Conference.* April http://www.eurocontrol.int/sites/default/files/content/documents/events/Presentations/130423-icns13-redeborn.pdf (accessed 8/23/13).

Riche, N., Duvinage, M., Mancas, M., Gosselin, B., & Dutoit, T. (2013). Saliency and Human Fixations: State-of-the-art and Study of Comparison Metrics. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*: Syndey, Australia.

RTCA (2011). *Safety, Performance and Interoperability Requirements Document for Airborne Spacing Flight Deck Interval Management (ASPA-FIM)*-DO-328. RTCA: Washington, DC.

Saaty, T. L. (2008). Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors – The Analytic Hierarchy/Network Process. *RACSAM (Review of the Royal Spanish Academy of Sciences, Series A, Mathematics) 102* (2): 251–318. Retrieved 2013-6-22.

Moacdieh, N. & Sarter, N. (2014). Display Clutter: A Review of Definitions and Measurement Techniques *Human Factors: The Journal of the Human Factors and Ergonomics Society,* (published online before print on July 3, 2014)*.*

Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal 27*(3): 379–423.

Shay, R., Swieringa, K., & Baxley, B. (2012). Aircraft Configuration and Flight Crew Compliance with Procedures While Conducting Flight deck based Interval Management (FIM) Operations. *12th AIAA Aviation Technology, Integration, & Operations (ATIO) Conference*: Indianapolis, IN.

Swieringa, K., Murdoch, J. L., Baxley, B., & Hubbs, C. (2011). Evaluation of an Airborne Spacing Concept, On-Board Spacing Tool, and Pilot Interface. *11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*: Virginia Beach, VA.

Tatler, B.W., Baddeley, R.J. & Gilchrist, I.D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research 45*, 643–659

Taylor, R.M., Selcon, S.J., & Swinden, A.D. (1995). Measurement of Situational Awareness and Performance: A Unitary SART Index Predicts Performance on a Simulated ATC Task, *Human Factors in Aviation Operations*, Edited by R. Fuller, N. Johnston and N. McDonald. Avebury Aviation, Ashgate Publishing, Aldershot, Hants

Tole, J., Stephens, A., Vivaudou, M., Harris, R. & Ephrath, A. (1982). entropy, Instrument Scan, and Pilot Workload. In *Proceedings of the IEEE Conference on Systems, Man & Cybernetics*: Seattle, WA.

Vu, K., Strybel, T.Z., Kraut, J., Bacon, P., Minakata, K., Nguyen, J., Rotterman, A., Battiste, V., & Johnson, W. (2010). Pilot and Controller Workload and Situation Awareness with Three Traffic Management Concepts. *29th Digital Avionics Systems Conference (DASC):* Salt Lake City, UT.

Wilson, S. R., Murdoch, J. L., Hubbs, C. E., & Swieringa, K. A., (2013). Evaluation of Flight deck-based Interval Management Crew Procedures Feasibility, *AIAA Modeling and Simulation Technologies Conference and Exhibit*: Boston, MA.

9   Appendices

The following appendices are included:

A - Display Positions & Formats

B - Aural Indication Specifications

C – Scenario Environment Details

D - Post-run Questionnaire Items *(not all items analyzed)*

E – Post-experiment Questionnaire Items *(not all items analyzed)*

9.1 Appendix A. IMSACE Display Positions & Formats

FAA (AC 25.1322-1) defines primary field of view in both horizontal and vertical directions. Optimal horizontal field of view is within +/-15 degrees of line of sight, with the maximum acceptable deviation from line of sight being +/-35 degrees. Optimal vertical field of view is +/-15 degrees off a nominal 15 degree downward line of sight, with the maximum acceptable deviation from this line of sight being 40 degrees upward and 20 degrees downward.

Horizontal angles expressed here are from the perspective of the left seat, but are mirrored for the right seat. Vertical angles are the same for both seats. Angles are expressed to the nearest degree and, unless otherwise mentioned refer to the center of the displays from the pilot eyepoint. In terms of horizontal angles, the PFD is located directly in the pilot's line of sight (0 degrees), displays located to the right of that instrument are the: AGD (22 degrees), the FMS MCDU (26 degrees) and the Upper EICAS display (32 degrees). To the left of the PFD, are the EFB in the foreward position (60 degrees), and the EFB in the aft position (73 degrees to the inner most top corner, and 90 degrees to the outermost top corner). As such, neither of the EFB positions was located in the primary field of view. In terms of vertical field of view, the AGD was located 34 degrees upward from horizontal and so 49 degrees upward from the nominal 15 degree downward gaze. The PFD and Upper EICAS were both downward angles of 32 degrees from the horizontal (17 from the nominal downward gaze). The CDU was 49.5 degrees down from horizontal (34.5 degrees from the nominal downward gaze). Both the EFB in the forward and in the aft positions were downward gazes of approximately 63 degrees from horizontal, or 48 degrees down from the nominal gaze direction, although the distance from the pilot eyepoint in the lateral direction for the EFB in the aft position was roughly twice that as in the fore position. As such, only the PFD, Upper EICAS , and CDU were within the primary field of view. EFB positions were chosen to reflect locations observed as implementations on this aircraft from review of literature and websites. The position of the AGD was selected to be similar to that chosen in UPS and MITRE studies.

The following pages in this Appendix show display formats for the different Avionics Conditions.
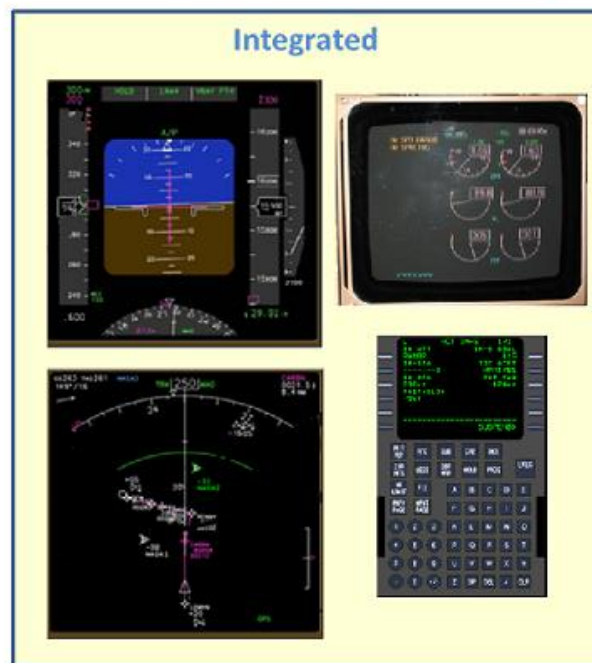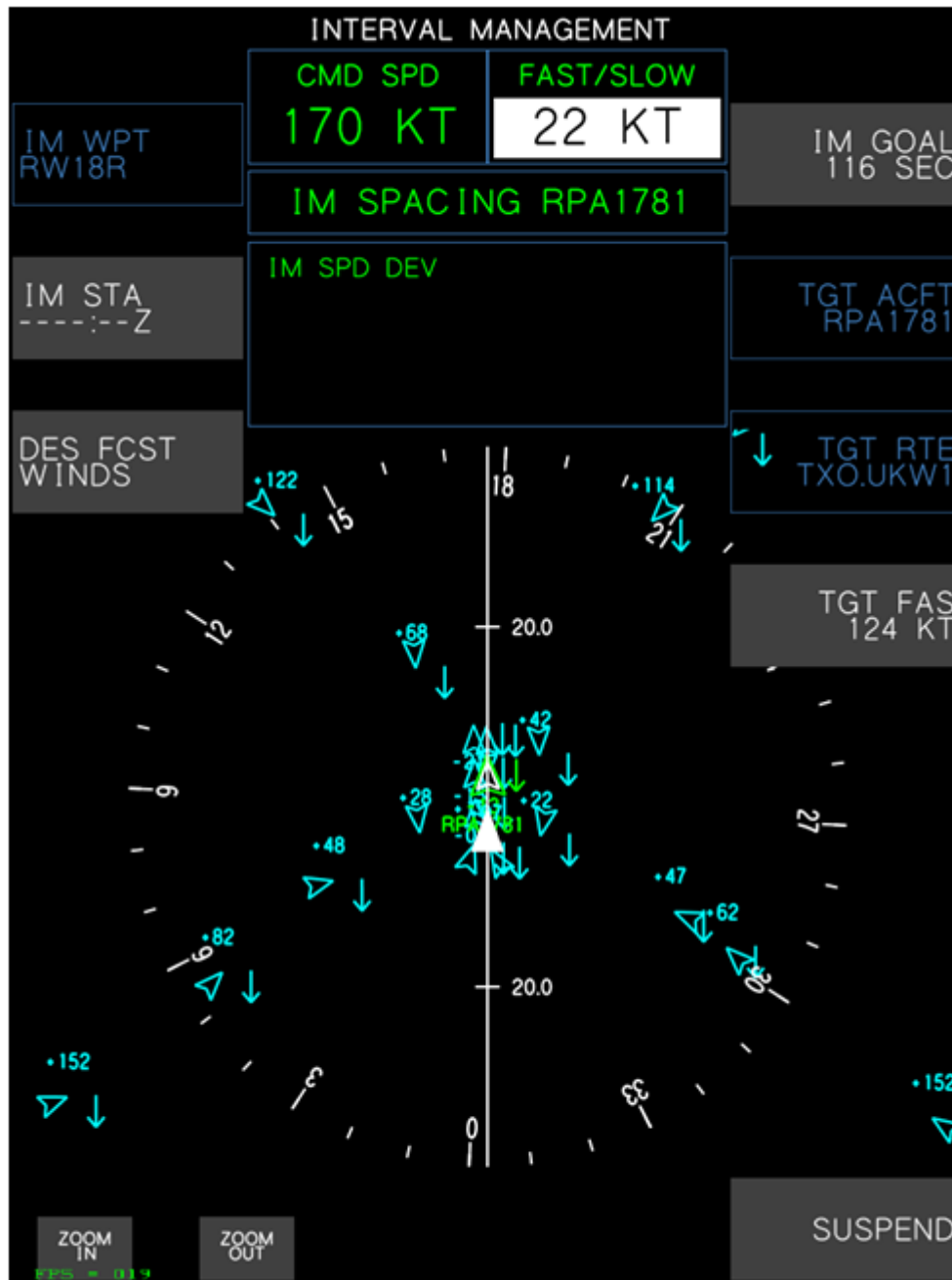
IMSACE Display Conditions

New Speed Target
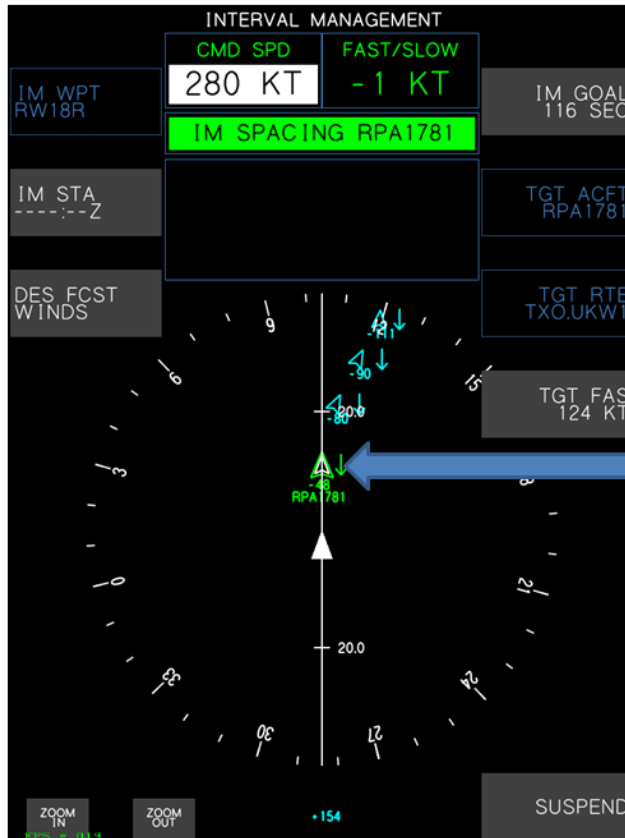
Reverse highlight white, until you dial the speed in.

If correct speed not in MCP 10sec after received, CMD SPD blinks as a "Reminder"

Target Aircraft

"SUSPEND" key if you'd want To discontinue IM. We'll ask you to continue the scenario, but it will help us understand your concerns later to know exactly when you thought conditions warranted this.
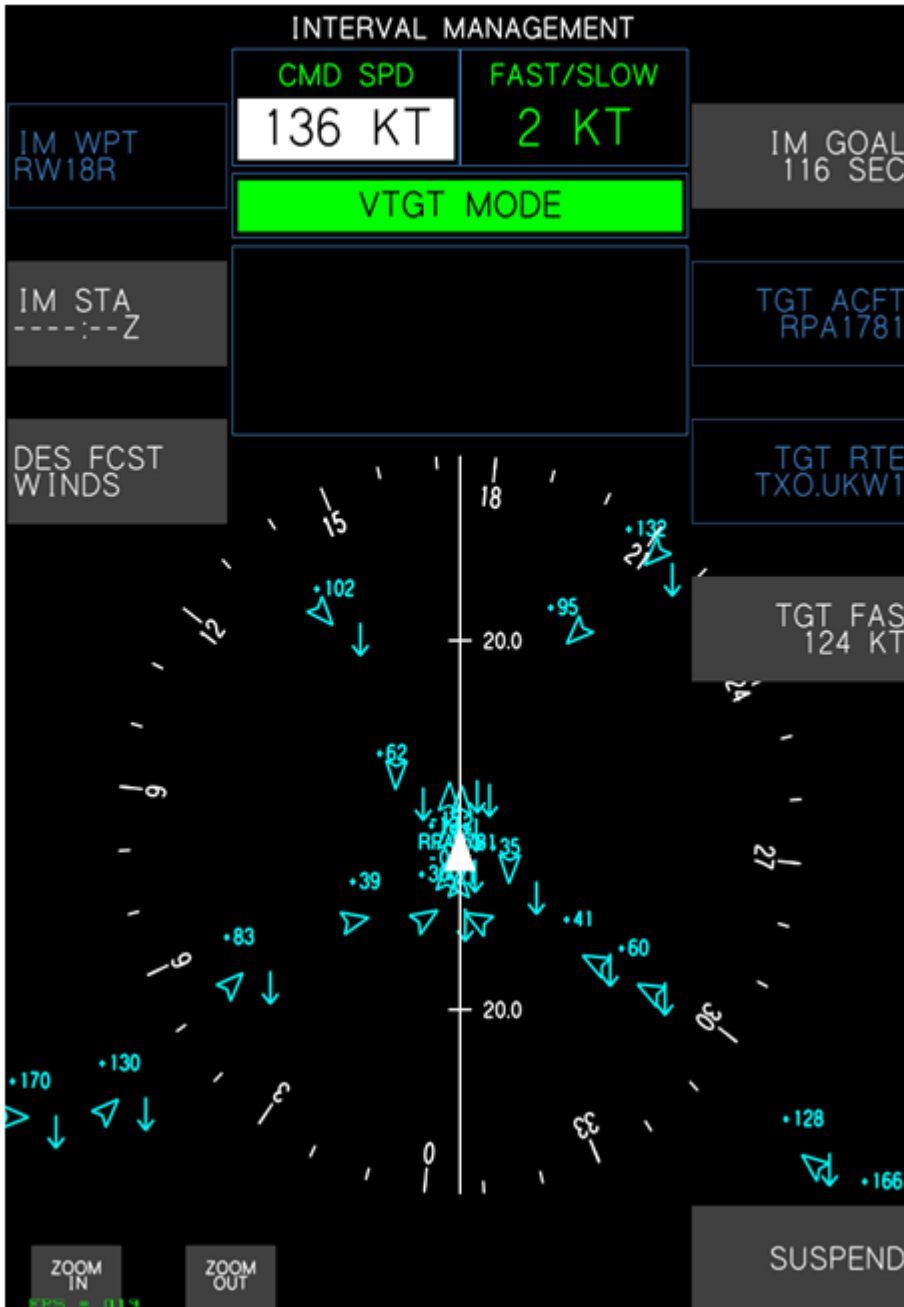
Deviation from Speed Profile – Conformance Deviation Indication

INTERVAL MANAGEMENT

| CMD SPD | FAST/SLOW |
| 170 KT | 22 KT |

IM WPT
RW18R

IM GOAL
116 SEC

IM SPACING RPA1781

IM SPD DEV

IM STA
----:--Z

TGT ACFT
RPA1781

DES FCST
WINDS

TGT RTE
TXO.UKW1

TGT FAS
124 KT

SUSPEND

ZOOM IN   ZOOM OUT

Conformance deviation Highlights white and blinks when 7 knots off for over 12 seconds and not converging.
22 KT means ownship is 22 knots fast. -11 KT would mean 11 knots slow.

Final speed provided
VTGT = VREF30+5

After AGL=1000'
Guidance Removed

INTERVAL MANAGEMENT

IM WPT
RW18R

--- ---

IM GOAL
116 SEC

IM STA
----:--Z

TGT ACFT
RPA1781

DES FCST
WINDS

TGT RTE
TXO.UKW1

TGT FAS
124 KT

ZOOM IN

ZOOM OUT

SUSPEND
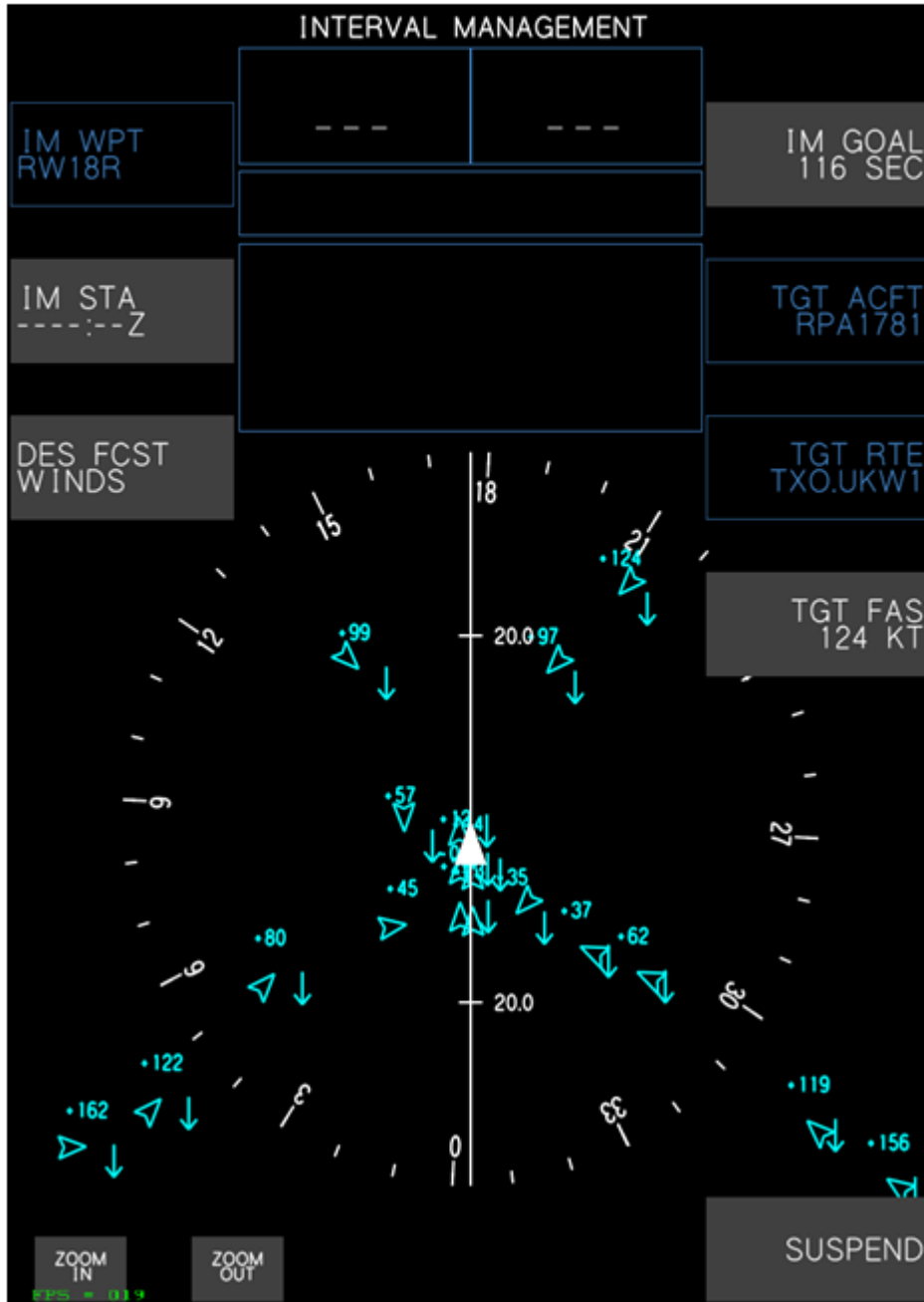
ADS-B GUIDANCE DISPLAY

- **New Speed** – New CMD SPD with white light
  - Turns off once it has been dialed it in to MCP

- **Reminder** - CMD SPD blinks if correct speed not dialed in by 10 seconds after occurrence.

- **Conformance Error** - Fast/Slow indicator blinks

28

INTEGRATED – PFD: Speed Target
Onset Speed Boxed

Boxed speed on Occurrence.

Box blinks after 10 seconds if correct speed not entered.

Profile Bug

CMD Spd

EICAS message

MCDU IM Page

SPD
230 IM

260
240
219
200
180

F

VNAV PTH   LNAV   CMD

20 — 20
10 — 10
10 — 10

DH200

IM SPD DEV
IM SPACING

TAT +11°C
1.55
1.28
1.0
1.5
2.0

L   ACT IM-S   1/1
IM WPT   IM-S GOAL
RW18R   116
IM-STA   TGT ACFT
----:--Z   RPA1781
IM SPD   TGT FAS
230KT   124KT
FAST/SLOW
-9KT
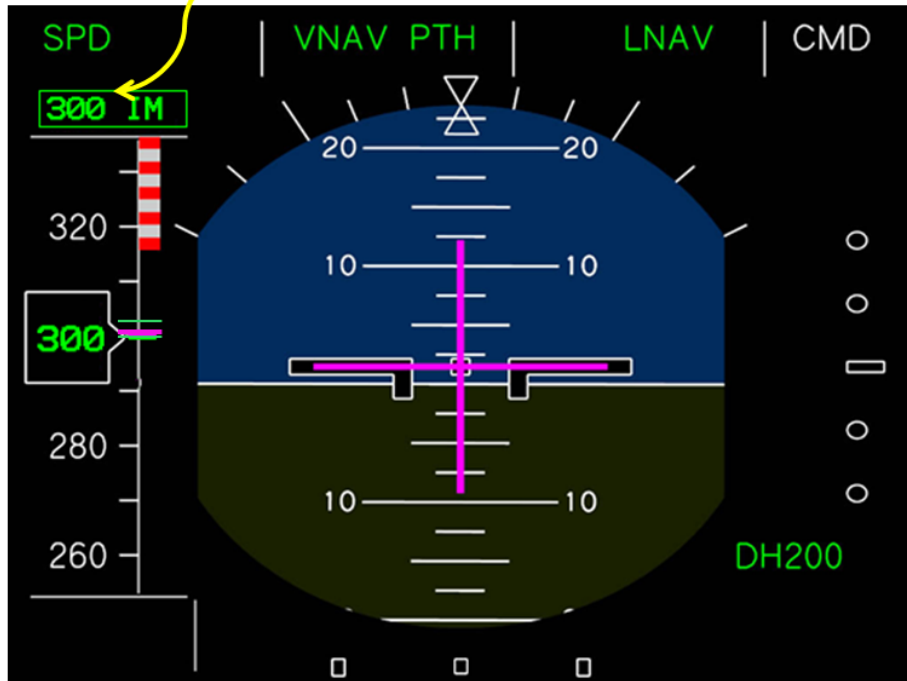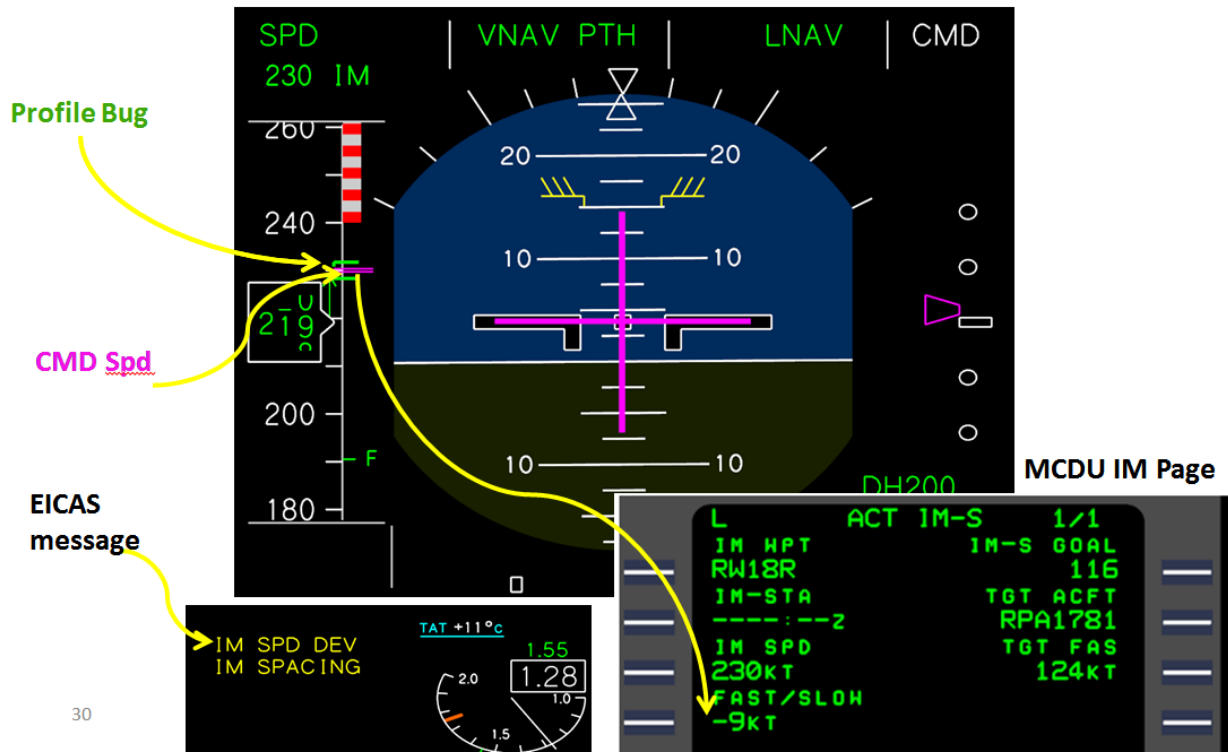
30

INTEGRATED: MCDU IM PAGE

- IM WPT (RWY)

- IM SPD (230 KT)

- FAST/SLOW (-9 KT)

- IM-S Goal (116 sec)

- Target Aircraft (RPA1781)

- Target FAS (124 kt)

- "SUSPEND"

INTEGRATED - Navigational Display
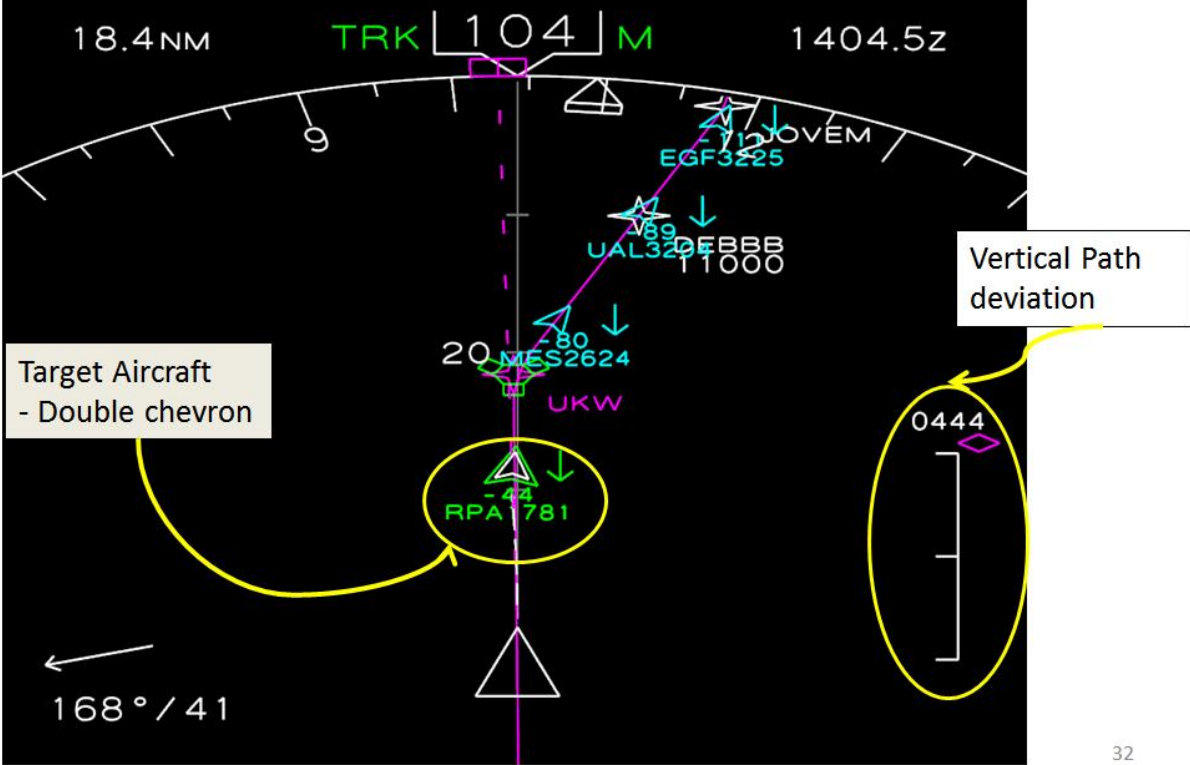
Integrated Condition for Last Speed Given
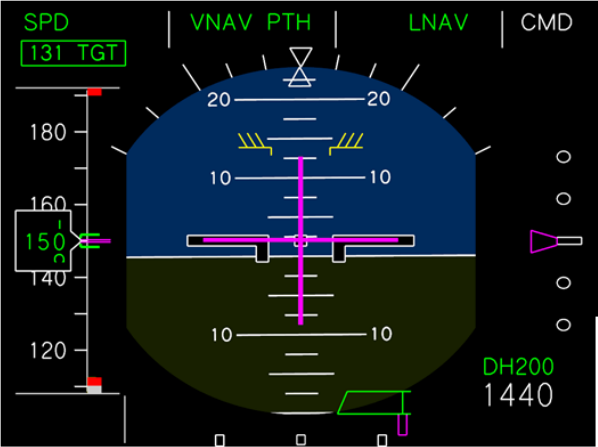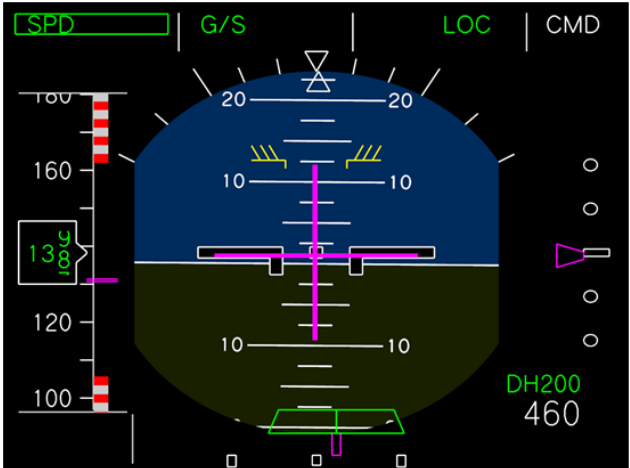
9.2   Appendix B. IMSACE Aural Indication Details.

The aural indication used consisted of two 300 Hz pulses, each 1050msec long with a 450msec inter-pulse interval.
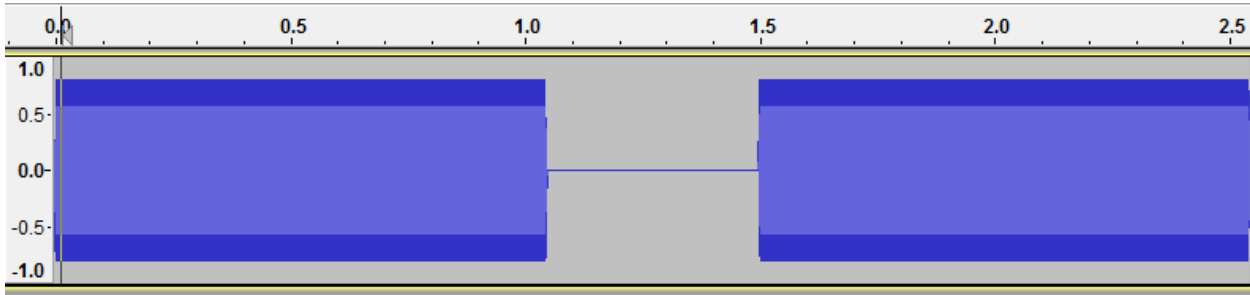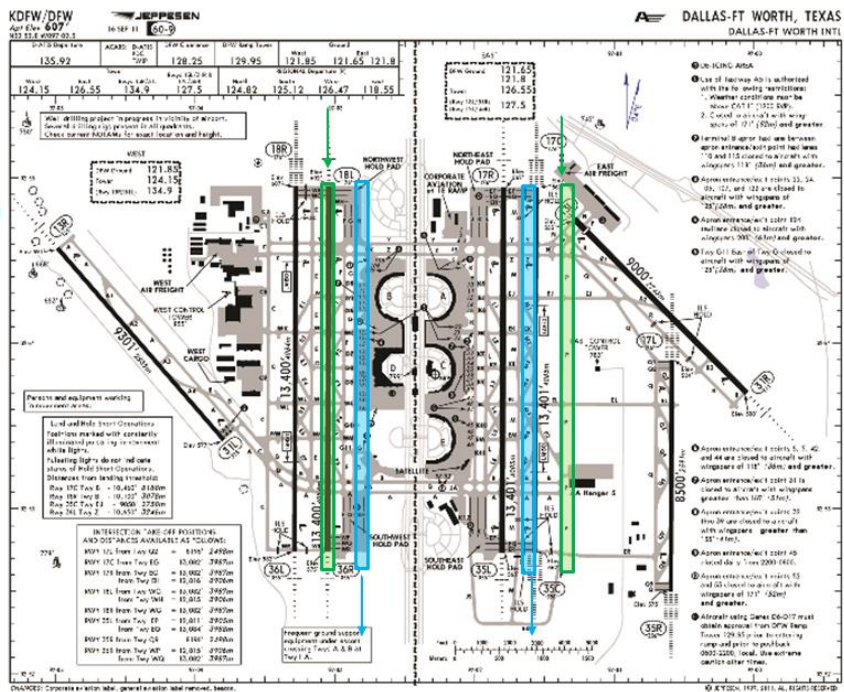

Each pulse was ramped up and down over a period of approximately 20msec.
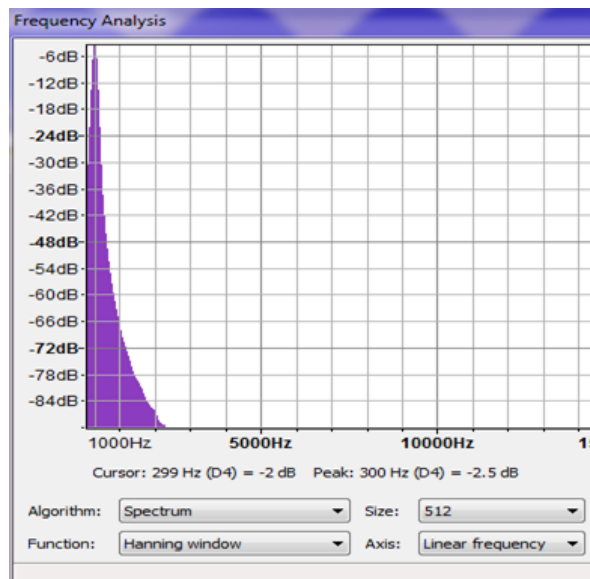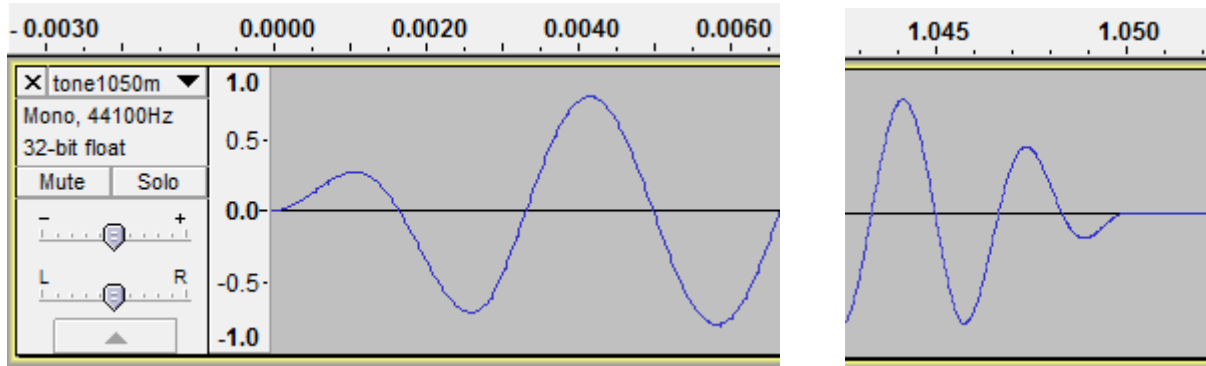
The peak of this distribution is 300Hz (D4, or "middle D")

9.3   Appendix C. Scenario Environment.
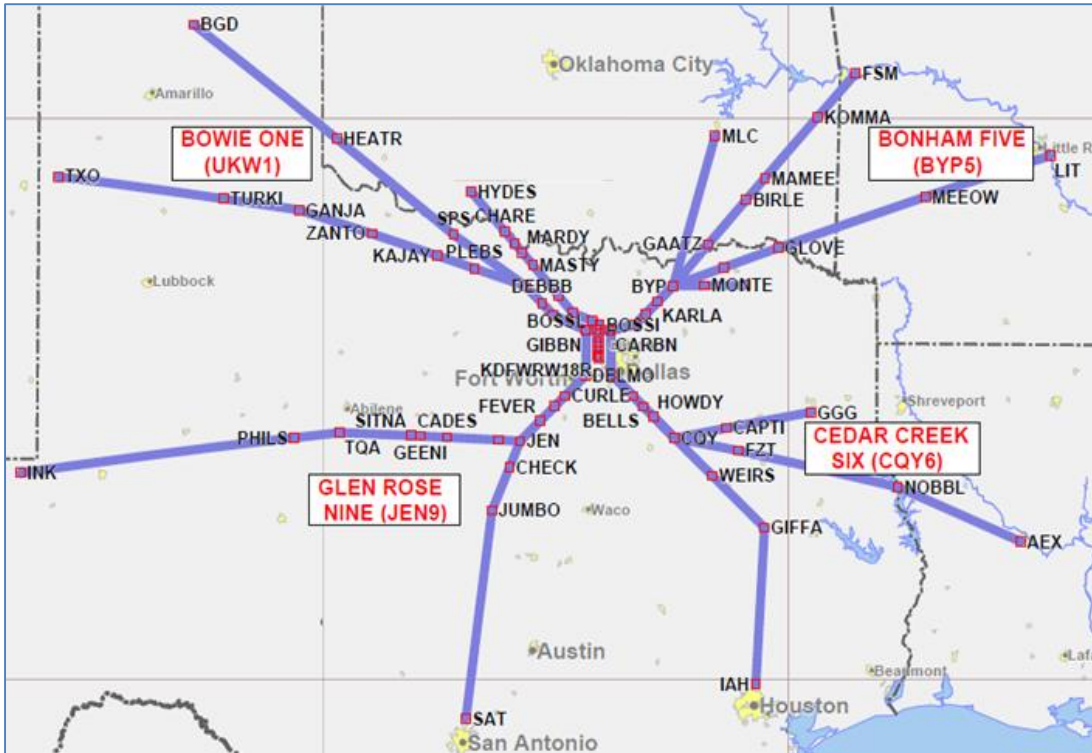
KDFW Arrivals to 17C and 18R

- 30 ops/hour/runway
  (120 ops/hr)
- **Arrivals on outside**
  (17C & 18R)
- **Departures on inside**
  (17R & 18L)

9.4 Appendix D – Post-Run Questionnaire Items

(all ratings are scored more positive for larger values)

IMSACE Post-Run Questionnaire   Crew#:____, ID#:____, Date: 11/   /12 Run#:___

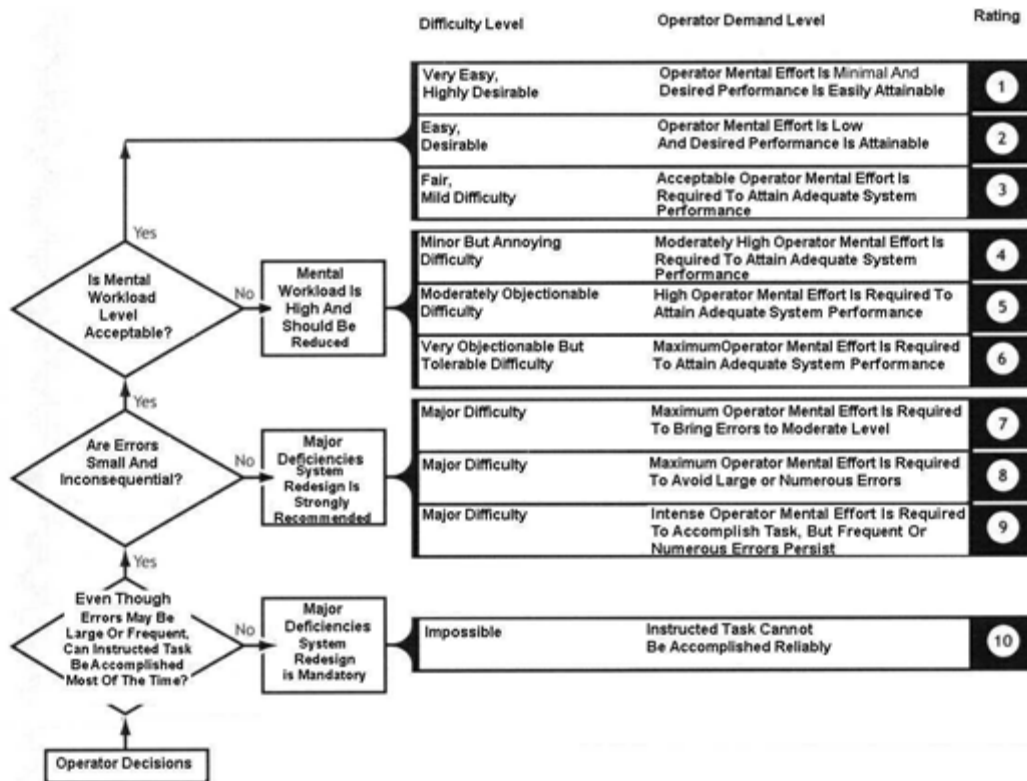**(Disregard) questions 1 & 2 for Training Runs)**

- In this run, you participated as: ☐ Pilot Flying   ☐ Pilot Monitoring/Not Flying

- In this run, you were in the: ☐ Left seat   ☐ Right Seat

Workload Ratings

3,4,5,6,7. Please use the scales to indicate your workload on the various facets described.

How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?	MENTAL DEMAND Low — High
How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?	TEMPORAL DEMAND Low — High
How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?	PERFORMANCE Good — Poor
How hard did you have to work (mentally and physically) to accomplish your level of performance?	EFFORT Low — High
How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?	FRUSTRATION Low — High

124

Workload Ratings (Modified Cooper-Harper)

Difficulty Level	Operator Demand Level	Rating
Very Easy, Highly Desirable	Operator Mental Effort Is Minimal And Desired Performance Is Easily Attainable	1
Easy, Desirable	Operator Mental Effort Is Low And Desired Performance Is Attainable	2
Fair, Mild Difficulty	Acceptable Operator Mental Effort Is Required To Attain Adequate System Performance	3
Minor But Annoying Difficulty	Moderately High Operator Mental Effort Is Required To Attain Adequate System Performance	4
Moderately Objectionable Difficulty	High Operator Mental Effort Is Required To Attain Adequate System Performance	5
Very Objectionable But Tolerable Difficulty	Maximum Operator Mental Effort Is Required To Attain Adequate System Performance	6
Major Difficulty	Maximum Operator Mental Effort Is Required To Bring Errors to Moderate Level	7
Major Difficulty	Maximum Operator Mental Effort Is Required To Avoid Large or Numerous Errors	8
Major Difficulty	Intense Operator Mental Effort Is Required To Accomplish Task, But Frequent Or Numerous Errors Persist	9
Impossible	Instructed Task Cannot Be Accomplished Reliably	10

Flow chart:

Is Mental Workload Level Acceptable? — No → Mental Workload Is High And Should Be Reduced

Are Errors Small And Inconsequential? — No → Major Deficiencies System Redesign Is Strongly Recommended

Even Though Errors May Be Large Or Frequent, Can Instructed Task Be Accomplished Most Of The Time? — No → Major Deficiencies System Redesign is Mandatory

Operator Decisions

8,9. Follow the flow chart above to characterize the workload **you** experienced in this scenario

- ➤ Rating of **your average** workload level: _____
- ➤ Rating of **your peak** workload level: _____

(10-11 deleted)

12. If your peak workload was over 3. Please indicate where and why. You can also use the space below for any additional comments regarding workload.

Situation Awareness and Crew Coordination Ratings

Respond to each of the statements shown below using a scale ranging from "1" (Low) to "7" (High). Circle one number in conjunction with each statement.

	Low						High
13. Demand on Attentional Resources: Rate your overall impression of the scenario in terms of how much attention and effort was required to successfully perform the tasks. Items to consider include: the likelihood of the situation changing suddenly, the degree of complexity associated with the scenario, and the number of variables changing in the scenario.	1	2	3	4	5	6	7
14. Supply of Attentional Resources: Rate the degree of spare attention that you had available to perform tasks other than your primary task of flying the aircraft. Items to consider include: the amount of focus and concentration needed to complete the tasks and how divided you attention was between flying the aircraft and other tasks.	1	2	3	4	5	6	7
15. Understanding: Rate your overall understanding of the events in the previous scenario. Items to consider include: the quantity and quality of information, and the familiarity you had with events in the scenario.	1	2	3	4	5	6	7

Ask me later about this run, regarding: _____

126

Acceptability

16-37.	Completely Disagree						Completely Agree
I was aware of commanded speed changes within an appropriate timeframe.	1	2	3	4	5	6	7
I was able to implement the speed changes in the MCP within an appropriate timeframe.	1	2	3	4	5	6	7
The commanded speeds were operationally acceptable and appropriate.	1	2	3	4	5	6	7
The frequency of IM speed commands was acceptable throughout the scenario.	1	2	3	4	5	6	7
The speeds given did not require an uncomfortable level of speedbrake use.	1	2	3	4	5	6	7
IM speeds and their timing are consistent with expectations for this type of arrival.	1	2	3	4	5	6	7
I maintained adequate awareness of my lead aircraft throughout the scenario.	1	2	3	4	5	6	7
I was never surprised to notice that the speed target had changed and wondered how long ago it had.	1	2	3	4	5	6	7
I was aware of speed profile conformance throughout the scenario.	1	2	3	4	5	6	7
The avionics and annunciations provided in this scenario were not overly distracting.	1	2	3	4	5	6	7
Use of the avionics and annunciations provided in this scenario are acceptable for conducting IM operations.	1	2	3	4	5	6	7

127

	Completely Disagree						Completely Agree
The avionics and annunciations provided in this scenario appropriately supported my ability to detect speed target onsets in a timely manner.	1	2	3	4	5	6	7
The avionics and annunciations provided in this scenario provided me appropriate support for detecting deviations from speed profile in a timely manner.	1	2	3	4	5	6	7
The events I experienced in this scenario are operationally realistic.	1	2	3	4	5	6	7
Time required for IM tasks did not detract from having appropriate SA for other aspects of flight	1	2	3	4	5	6	7
In this scenario, I thought commanded speeds were safe and appropriate to fly.	1	2	3	4	5	6	7
I received IM speeds when in the process of completing other critical tasks.	1	2	3	4	5	6	7
At no point in this scenario did I feel that that the commanded speed conflicted with other available information.	1	2	3	4	5	6	7
At all times in this scenario, I felt comfortable flying the commanded speeds.	1	2	3	4	5	6	7
I did not find the spacing tool frustrating during this scenario.	1	2	3	4	5	6	7
I was comfortable with the location of my target aircraft throughout the run	1	2	3	4	5	6	7
Apart from IM operations, this scenario was typical in terms of workload for an arrival at a busy airport.	1	2	3	4	5	6	7

9.5 Appendix E. Post-Experiment Questionnaire Items.

(Some data were selected out for the current analyses.)

IMSACE Post-Experiment Questionnaire   Crew#:___, ID#:___, Date:_____

Please put a check in a box to indicate, for each pair of avionics conditions, which you prefer *in terms of providing you the best support in detecting speed onsets in a timely manner and allowing you to maintain adherence to the speed profile.*

EFB in Aft Position (no AGD)   EFB in Forward Position (no AGD)

Strongly prefer				No difference				Strongly prefer

EFB in Aft Position (no AGD)   EFB in Aft Position + AGD

Strongly prefer				No difference				Strongly prefer

EFB in Aft Position (no AGD)   Integrated Displays (PFD, ND, CDU)

Strongly prefer				No difference				Strongly prefer

EFB in Forward Position (no AGD)   EFB in Aft Position + AGD

Strongly prefer				No difference				Strongly prefer

EFB in Forward Position (no AGD)   Integrated Displays (PFD, ND, CDU)

Strongly prefer				No difference				Strongly prefer

EFB in Aft Position + AGD   Integrated Displays (PFD, ND, CDU)

Strongly prefer				No difference				Strongly prefer

1

Please indicate the degree to which you consider these solutions operationally acceptable (i.e., in terms of operational risk, heads down time, workload, total situation awareness, scan pattern):

It is acceptable to conduct IM operations with the Forward EFB alone

Strongly agree				Borderline				Strongly disagree

It is acceptable to conduct IM operations with the Aft EFB alone

Strongly agree				Borderline				Strongly disagree

It is acceptable to conduct IM operations with the Aft EFB with the AGD

Strongly agree				Borderline				Strongly disagree

It is acceptable to conduct IM operations with the Integrated Displays

Strongly agree				Borderline				Strongly disagree

Please consider the tradeoff of the following aural indications, and use the scale to rate your impression of how they did/would affect performance.

To ensure timely detection, an aural indication of speed target onset is/would be:

Unnecessary							Necessary
Too Distracting							Appropriate

To ensure timely detection, an aural indication of un-entered speed targets is/would be:

Unnecessary							Necessary
Too Distracting							Appropriate

To minimize deviations from the speed profile, an aural indication of speed target onset is/would be:

Unnecessary							Necessary
Too Distracting							Appropriate

2

130

Please tell us more about how you think these conditions should be brought to your attention:

Target Onset:

Missed Speed Entry:

Conformance Deviation:

Annunciation Conditions

An indication was provided if you did not enter a newly presented target speed within 10 seconds.

Did you receive any of these indications?   ☐YES   ☐NO

If you did, how frequently would you consider these to have been false alarms, i.e., you had entered the speed as soon as you noticed it.

Almost Never				50/50				Almost all

What do you think is a reasonable condition under which a pilot should be reminded of a pending speed target? _____
_____

An indication was provided if you were greater or less than the calculated profile speed by over 7 knots for more than 12 seconds, and not converging to the profile speed.

Did you receive any of these indications?   ☐YES   ☐NO

If you did, how frequently would you consider these to have been false alarms, i.e., you had initiated the speed change as soon as practical, and in a manner consistent with passenger comfort, and the airplane would not respond fast enough to avoid triggering this condition.

Almost Never				50/50				Almost all

3

131

What do you think is a reasonable speed deviation to allow without annunciation, for how long? _____
_____

How would you rate the overall effect of IM operations on your workload, as compared to normal operations in a similar aircraft under similar conditions?

Greatly reduces workload			No effect			Greatly increases workload

How would you rate the overall effect of IM operations on your situation awareness of arrival speeds, as compared to normal operations in a similar aircraft under similar conditions?

Greatly reduces SA			No effect			Greatly increases SA

How would you rate the overall effect of IM operations on situation awareness of other elements, as compared to normal operations in a similar aircraft under similar conditions?

Greatly reduces SA			No effect			Greatly increases SA

Flying commanded speeds via the auto-throttle / MCP was acceptable

Strongly Agree			Borderline			Strongly Disagree

Would you expect any new phraseology to accompany during M&S operations? Please explain:

_____

_____

What advantages/disadvantages (operational risks) do you see with Interval Management operations?

_____

_____

4

How well did your standard crew coordination and cockpit resource management strategies map onto the use of IM speed targets to achieve a spacing constraint?

Extremely poorly – specific training required							Extremely well –no adaptation required

What cockpit resource management / crew coordination procedures would you recommend?

_____

_____

What type of impact did Interval Management procedures have on your primary tasks?

As PF: _____

As PM: _____

Validity and Effectiveness of the Study

How would you rate the training you received?

Extremely poor							Extremely good

Please tell us how we can improve: _____

_____

How realistic was the simulated cockpit (displays, control panel, etc)

Completely unrealistic							Completely realistic

Please tell us how we can improve: _____

_____

5

133

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* 01-08-2015	2. REPORT TYPE Technical Publication	3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE

Avionics Configuration Assessment for Flightdeck Interval Management: A Comparison of Avionics and Notification Methods

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Latorella, Kara A.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
411931.02.01.07.13.03

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER

L-20342

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)
NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NASA/TP-2015-218788

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 03
Availability: STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Flightdeck Interval Management is one of the NextGen operational concepts that FAA is sponsoring to realize requisite National Airspace System (NAS) efficiencies. Interval Management will reduce variability in temporal deviations at a position, and thereby reduce buffers typically applied by controllers – resulting in higher arrival rates, and more efficient operations. Ground software generates a strategic schedule of aircraft pairs. Air Traffic Control (ATC) provides an IM clearance with the IM spacing objective (i.e., the TTF, and at which point to achieve the appropriate spacing from this aircraft) to the IM aircraft. Pilots must dial FIM speeds into the speed window on the Mode Control Panel in a timely manner, and attend to deviations between actual speed and the instantaneous FIM profile speed. Here, the crew is assumed to be operating the aircraft with autothrottles on, with autopilot engaged, and the auto-flight system in Vertical Navigation (VNAV) and Lateral Navigation (LNAV); and is responsible for safely flying the aircraft while maintaining situation awareness of their ability to follow FIM speed commands and to achieve the FIM spacing goal.

15. SUBJECT TERMS

Air traffic control; Flightdeck; Interval management; National airspace system; Spacing

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON STI Help Desk (email: help@sti.nasa.gov)
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	149	19b. TELEPHONE NUMBER *(Include area code)*
U	U	U			(757) 864-9658

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18