

# PREFERENCE AND OPERATIONAL ACCEPTABILITY OF FLIGHTDECK INTERVAL MANAGEMENT AVIONICS

*Kara Latorella & Raleigh B. Perry, NASA Langley Research Center, Hampton, VA*

*Richard Shay, National Institute of Aviation, Hampton, VA*

*William Merritt & Robert Cameron, Northrup Grumman Technical Services, Hampton, VA*

## Abstract

This study investigated the relative acceptance of different avionics implementations that present Flightdeck Interval Management (FIM) speeds and speed deviations to commercial pilots, and for indications of conditions that require action. Results indicate a clear preference for an Avionics condition where target speed information was provided in the primary flight display, relevant traffic information was provided in the navigational display, IM clearance information and conformance information was provided in the Multi-function Control Display Unit (MCDU), and the engine-indicating and crew-alerting system (EICAS) display showed conformance deviation alerts; and a condition in which all this information was presented only in an Electronic Flight Bag (EFB)-like display (with extended functionality) implemented under the side window for both pilots. Other Avionics conditions tested were less desirable. These were conditions in which the EFB was mounted aft of, and below the side window – with and without an auxiliary display that repeated speed target and conformance information in the primary field of view. Results also indicate a preference for aural indications to direct attention to new speed targets, as a reminder to enter these when not done in a timely manner, and to convey when the aircraft has deviated significantly from the calculated FIM speed profile. The aural indications, thresholds for reminders and conformance indications used in this study were found to be appropriate. In general FIM, as implemented in this study, was perceived as having no deleterious effect on workload or crew coordination; and, under some conditions, was reported to have improved situation awareness of arrival speeds and general conditions during approach and descent.

## Introduction

### *Interval Management*

Flightdeck Interval Management (FIM) operations allow an aircraft (hereafter referred to as the *IM aircraft*) to space off another aircraft (hereafter referred to as the *Target To Follow* (TTF)), using speed guidance. During IM operations, Air Traffic Control (ATC) provides an IM clearance to the IM aircraft which includes: the spacing interval to achieve, the TTF identifier, and the point along the path at which to achieve the spacing goal. Pilots of the IM aircraft enter the clearance information in an onboard system containing the FIM algorithm. Once the TTF's Automatic Dependent Surveillance-Broadcast (ADS-B) signal is available, the IM aircraft can begin spacing. The algorithm uses a speed profile based on the standard terminal arrival route (STAR) in use, conforms to standard speed constraints in the terminal environment (*e.g.*, less than 250 knots indicated airspeed under 10,000 feet), and is adaptive to forecasted 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 [1], the crew is assumed to be operating the aircraft with autothrottles on and with autopilot engaged and the auto-flight system in Vertical Navigation (VNAV) and Lateral Navigation (LNAV). To support FIM, 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.

## Objective

The objective of this investigation was to assess different FIM Avionics configurations. We collected objective data (flightpath and speed profile deviations, and response times), oculometer data, and subjective data. This paper discusses the subjective ratings associated with the different Avionics conditions assessed as derived from Post-Experiment questionnaire data.

## Methods

This study was conducted as a Human-in-the-Loop (HITL) investigation in a realistic commercial flight deck simulation environment, with commercial pilot subjects.

### *Participants*

This study involved 12 crews of two pilots. Pilots were type rated in the same class as the simulated aircraft. Pilots comprising a crew were obtained from the same airline, and with the exception of one crew, were all constructed of one Captain and one First Officer. Eleven of these subjects reported demographic data. These pilots were generally experienced (between 19 and 40 years of flying, with the mean and median of this data being 28.9 and 30 years, respectively; with between 400 and 23600 commercial flight hours, with the mean and median of these data being 7115 and 4670 hours, respectively).

### *Apparatus & Scenarios*

The simulation environment was the NASA Langley Integration Flightdeck (IFD) simulator, which is similar to a Boeing 757 aircraft. Figure 1 shows the position of the EFB in the aft position, and the location of the ADS-B Guidance Display (AGD). Displays were implemented in the same positions for both pilots.

Scenarios required crews to fly from approximately 25,000 feet to land at Dallas-Fort Worth International Airport (KDFW). Scenarios began in level flight prior to top of descent, in VNAV Path autoflight mode engaged. subjects were given their IM clearance verbally and in written form.



**Figure 1. The Integration Flightdeck**

In this study, subjects 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 were 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 mode control panel (MCP) speed window open, until the flaps were extended, and the autoflight mode reverted to VNAV Path. Subjects 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 five knots to enable stabilized approach by 1000 feet above ground level. Scenarios concluded typically after roll out on touchdown, but occasionally in advance of that in order to save time (always after aircraft configuration for a stabilized approach was complete). Subjects were instructed to fly as they typically would, as though they had passengers in the back of the airplane, to respond to speed targets in a timely manner, to try to maintain speed conformance within seven knots, and to remain within 400 feet of the VNAV path.

Confederate Air Traffic Controllers provided realistic communications to both the simulator and to roughly 20 simulated aircraft in the environment. Prerecorded Automatic Terminal Information Service (ATIS) messages were available on the appropriate frequency.

## ***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).

Each crew evaluated four Avionics conditions: (1) Integrated – in which the FIM target speed was presented in the upper left corner of the primary flight display (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 in 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 provided 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 the aircraft current speed was more than seven knots different from the instantaneous speed on the FIM speed profile, 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 tone only if pilots significantly deviated from the speed profile.

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).

## ***Experimental Design***

The experiment had three factors: Avionics condition (with four levels: Integrated, EFB-Aft, EFB-Fore, EFB-AGD); Notification Method (with three levels: VVV, AAA, VAV); and Crew Role (with two levels: Pilot Flying, Pilot Monitoring). Avionics condition and Crew Role were within-crew variables. Notification Method was a between-crew variable. Order of Avionics conditions were counterbalanced over crews, and the assignment of scenarios to Avionics conditions was also counterbalanced.

## ***Dependent Variables***

Post-experiment questionnaire items asked subjects 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).

## Results

### Avionics and Notification Preference Ratings

Subjects 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,” with “No Difference” as the central response cue. Preference ratings were derived from these data according to the Analytical Hierarchy Process [2]. Normality was not demonstrated for all subset conditions (Shapiro-Wilk tests all  $p < 0.05$ ), however all Levene’s tests gave confidence in the equal variance assumption (all  $p > 0.05$ ). A mixed model testing Avionics, with Notification as a nested variable within Avionics on arcsin(square-root) transformed data show a significant effect of Avionics ( $F(3,63) = 154.945$ ,  $p < 0.001$ ). The effect of Notification Type did not rise to standard levels of significance ( $F(2,41.288) = 1.717$ ,  $p = 0.123$ ). Subsequent pairwise analyses (Sidak-adjusted to control Type 1 error) on Avionics conditions show significant differences (all  $p < 0.01$ ) among all pairs except between the EFB-Fore and the EFB-Aft+AGD conditions ( $p = 0.985$ ). Non-parametric assessments confirmed the Avionics effect for each Notification level separately (Friedman test, all  $p < 0.01$ ), and all showed the same pattern in mean ranks as are shown by means in Figure 2.

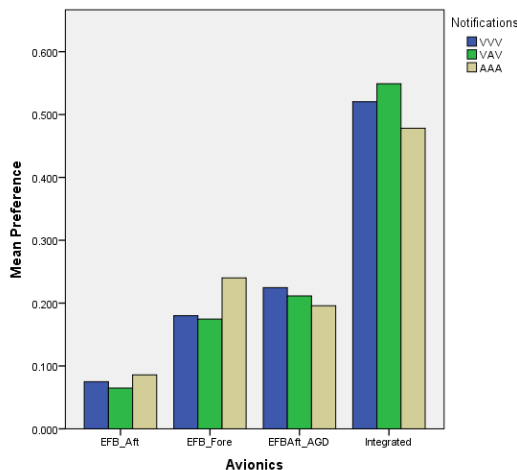


Figure 2. Preference Ratings for Avionics

### Avionics & Notification Acceptability Ratings

In the Post-Experiment questionnaire, subjects were asked to rate 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 registering a value of 1-9 describing their degree of agreement (“Strongly Disagree” to “Strongly Agree”) with the statement, “It is acceptable to conduct FIM operations with (each type of Avionics).” Ratings were scored such that a rating of one indicates strong disagreement that the Avionics condition rated acceptably supports FIM operations, and a rating of nine indicates strong agreement. These data failed normality and equal variance assumptions (Shapiro-Wilk tests and Levene tests’ all  $p < 0.05$ ). Friedman analyses were conducted to determine if Avionics condition ratings differed, for each Notification method separately. When Notifications included aural indications, Avionics conditions differed significantly (for AAA ( $X^2(3) = 16.158$ ,  $p = 0.001$ ), and less strongly for VAV ( $X^2(3) = 11.292$ ,  $p = 0.01$ )), but not when only visual indications were provided ( $X^2(3) = 4.477$ ,  $p = 0.214$ ) (Figure 3). For each Notification type, Wilcoxon pairwise comparisons for all pairs of Avionics conditions failed to reach significance for  $\alpha = 0.05$  Sidak-adjusted to 0.009. Binomial tests were used based on a cut point of five to test use of the scales. Subjects in both the VAV and AAA notification methods were more likely to rate the EFB-Fore and Integrated Avionics conditions as operationally acceptable (all  $p = 0.07$ ).

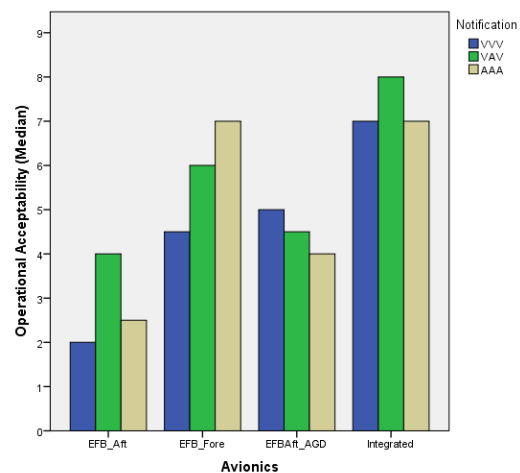


Figure 3. Acceptability Ratings for Avionics

### ***Comments on Avionics Conditions***

In addition to explicit questions regarding preference, subjects were asked to provide comments on their relative impressions of the different Avionics conditions. Twenty of the subjects provided positive commentary on the Integrated condition, and sixteen commented positively on the Fore-EFB condition. When subjects were critical of the Integrated condition, it was only to suggest that the font size indicating speed targets was too small, and that the flashing reminders did not persist for long enough. Those critical of the Fore-EFB 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 and form in turbulence conditions. With this condition, but more so with the Aft-EFB condition, subjects noted a concern for inducing vertigo. Subjects who commented on the AGD routinely mentioned that they overlooked changes to speed targets 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 Aft-EFB condition was useable.

Comments regarding the best way to direct attention to FIM events seemed to focus on the use of aural indications. While all subjects made mention of whether they would or would not want aural to augment visual indications, most subjects did not comment on the type of visual notifications they desired. In exception to this, several subjects did note the appropriate use of flashing to indicate FIM events. Four subjects suggested that this formatting method be used for all FIM events, three for onset only, three for reminders only, and one for onsets and reminders. Three subjects indicated changing colors on displays was effective for reminders and conformance, and one in particular suggested that reminders and conformance deviations be presented in red.

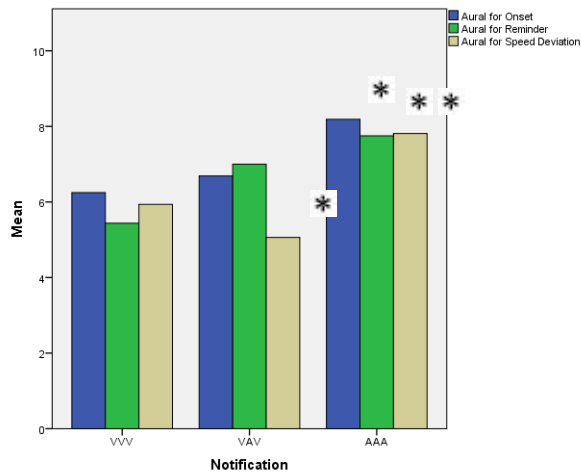
### ***Use of Aural in Notifications***

Subjects were asked about the necessity of, and distracting nature of, aural tones to indicate when a new speed target occurs, and when a reminder or speed deviation occurs. While two separate scales

were provided in the questionnaire, several subjects interpreted these as one scale and provided only one response. Ratings for responses on these scales differed only in seven instances (of 72 opportunities); for all these, ratings on the scale for distraction had higher scores than the un/necessary scale. In all but one case, ratings were no more than one scale point apart and in the remaining case, they differed by two points. It seemed that subjects' 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 one indicated that the aural indication would be considered unnecessary/distracting, whereas a score of nine indicates that the aural indication was considered necessary/appropriate.

All measures failed Shapiro-Wilk tests for normality (all  $p < 0.05$ ). Levene's tests failed to reach significance for ratings of aural for reminders and for conformance deviations (all  $p > 0.05$ ), but suggested unequal variances for ratings associated with ratings of aural for speed target onset. Notification type did not appear to be a significant factor in distinguishing ratings associated with aural for speed target onset (Kruskal-Wallis  $X^2(2) = 2.687014$ ,  $p = 0.261$ ) or for reminders ( $F(2,21) = 2.161$ ,  $p = 0.140$ ), but did show weak explanatory power for differences in ratings associated with speed conformance deviations ( $F(2,21) = 3.031$ ,  $p = 0.07$ ). Pairwise Tukey post hoc comparisons of Notification conditions for the conformance measure indicate no strongly significant findings, but show only weak evidence that subjects who experienced aural in all opportunities (onset, reminder, and conformance deviation) rated aural more positively than those who experienced only aural indications for conformance deviations ( $p = 0.063$ ). Assessing ratings for each Notification type separately shows no evidence of significant differences (Kruskal-Wallis, all  $p > 0.304$ ). Binomial tests with a cut-point at the scale midpoint implies that subjects in the AAA group (those who experienced aural indications for all three events) showed a significant tendency to use the more positive end of the scale when rating the acceptability of aural indications for speed target onset ( $p = 0.008$ ) and conformance deviations ( $p = 0.008$ ). Weak evidence indicates a tendency for subjects in both the AAA and VAV conditions to also positively rate the use of aural for reminders (both

$p=0.07$ ). It is important to note that for the VVV condition, these ratings were based on a hypothetical aural indication, whereas for the VAV and AAA condition, subjects had experienced aural indications. Figure 4 displays averaged preference ratings, (\*) indicates conditions in which subjects actually experienced aural indications.



**Figure 4. Preference for Aural Indications**

Those subjects who received aural indications were asked to rate whether the tone for these was urgent (1=not at all, 9=extremely) and annoying (1=not at all, 9=extremely). Subjects in the VAV condition tended to rate the tone as more urgent than those in the AAA condition (Kruskal-Wallis  $X^2(1)=6.792$ ,  $p=0.009$ ), but these groups did not differ on their ratings of how annoying the tone was (Kruskal-Wallis  $X^2(1)=0.433$ ,  $p=0.511$ ). Binomial tests with a cut point at the midpoint of these scales shows that, over all, subjects who received the aural indications were balanced in responses to whether the tone was more or less urgent than “borderline,” with all responses having central tendency (min=3, max=8) ( $p=0.754$ ). Binomial tests for ratings of tone annoyance showed a significant bias ( $p=0.021$ ), with more responses indicating that the tone was not significantly annoying.

### Comments on Notifications

Subjects were asked how the three types of FIM events (new speed target, conformance deviation, reminder) should be brought to their attention. Almost half of the subjects indicated a preference for aural indications (mentioning either tones or verbals)

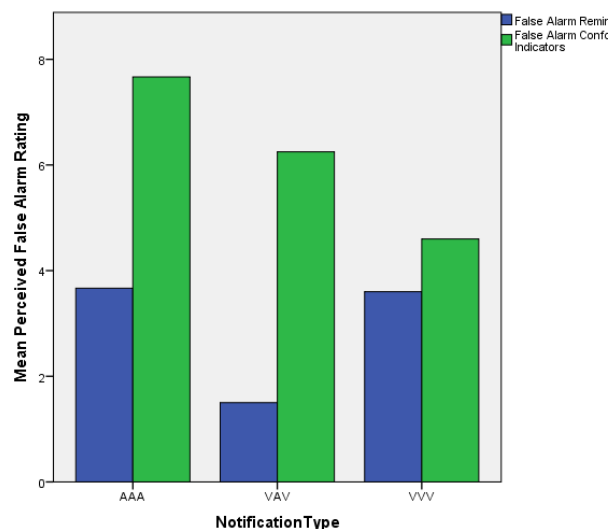
for all three of these events (10 of the 24 subjects). Two or three subjects indicated a preference for using aural indications for some subset of these three conditions (indication to dial in a speed target (onset or reminder) (3), indication of an error (reminder or conformance) (2), or only for onsets (3) or reminders (2)). Only two subjects indicated they’d desire only visual methods for notifying them of FIM events. Three subjects mentioned a preference for verbal messages (*i.e.*, “Interval speed change” for onsets, “Deviation Interval Speed” for conformance deviations). Those subjects who specified use of aural indications for all types of FIM events included five of the eight subjects who only had visual indications, and one subject who only had aural indications for conformance deviations. The remaining three subjects in the VVV condition preferred aural for onset of speed targets, and one of these also wanted an aural indication for conformance deviations. Of the seven remaining subjects in the VAV condition, six requested an aural for some FIM event (onset only (1), reminder only (2), reminder and conformance (2), onset and reminder (1)), and only one suggested that visual indications sufficed. Six of the eight subjects receiving aural indications for all events (Notification type AAA), suggested that this notification method was appropriate; two would have preferred use of aural only for onset indications, and one suggested that reminders be presented only visually. One subject indicated that there should be a different tone for each different FIM event, and four suggested that the tone indicating a new speed target onset should be distinct from what might be considered “errors” (reminders and conformance deviations).

### Reminder & Conformance Thresholds

All crews received both reminders and conformance deviation indications. However, whereas only half of the subjects recalled receiving reminder indications, all subjects indicated that they’d 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 one indicated “Almost Never,” a five indicated “50/50,” and a nine indicated “Almost All.” Shapiro-Wilk tests indicate normal distributions for both measures, for all but one



Notification subgroup (all  $p > 0.224$ ). The VAV group's ratings of reminder thresholds were suspected to be non-normal ( $p = 0.024$ ). Levene's tests provided confidence that both measures met the equal variance assumption for all groups (all  $p > 0.065$ ). ANOVAs show no significant effect of Notification type on these ratings for either the reminders ( $F(2,9) = 1.026$ ,  $p = 0.397$ ) or the speed conformance deviation indications ( $F(2,20) = 0.954$ ,  $p = 0.402$ ). For those subjects who provided ratings for both events, a paired t-test was conducted to ascertain if the false alarm experience for reminders different from that of conformance deviation indicators. Results indicate that these subjects found conformance deviation indications to be significantly more likely, on average, to be characterized as a false alarm than reminders ( $t(11) = -3.023$ ,  $p = 0.012$ ). Figure 5 shows that means of likelihood-to-be-perceived-as-false-alarms are higher for conformance deviations than for reminders.



**Figure 5. Perceived False Alarms**

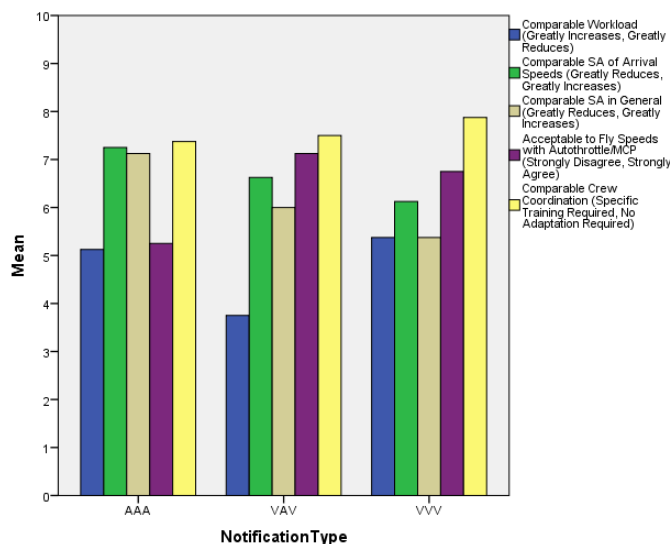
### Comments on FIM Event Thresholds

Subjects were asked to comment on the logic associated with receiving reminders and conformance deviation indications. Most (11) of those who commented on the reminder logic (14) agreed with the implementation tested, notifying the crew if the speed target had not been dialed into the MCP within 10 seconds. Two subjects 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. Subjects did not always express their preferences in terms of all these parameters. Of the 17 who indicated a tolerance in terms of a delta speed, responses included: five knots (by two subjects), seven knots (8), 10 knots (5), 12 knots (1), and 15 knots (1). Of the 15 subjects provided an out-of-tolerance time, six indicated 10 seconds out of tolerance (with speed tolerances ranging from five knots to 12 knots); three indicated 12 seconds out of tolerance (with speed tolerances all at seven knots); three indicated 15 seconds out of tolerance (with only one subject reporting a speed tolerance, this at 10 knots); and two subjects suggested an out of tolerance time of 20 seconds (one subject associating this with a seven knot speed tolerance).

### Operational Impact

Subjects were asked several questions regarding the operational impact of flying Interval Management operations, and how these compared to current operations. The specific questions follow with the cues used for anchors on these nine-point scales: "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); "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 awareness (SA), Greatly increases SA); "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); "Flying commanded speeds via the auto-throttle / MCP was acceptable" (Strongly Disagree, Strongly Agree); "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). Figure 6 shows the mean responses on these scales.



**Figure 6. Ratings of Operational Impact**

Shapiro-Wilk tests show some significant departures from normality over levels of Notification (where  $p < 0.05$ ), however Levene tests assure equality of variances (all  $p > 0.078$ ). Parametric statistics, robust to departures from normality, were selected to analyze how responses were affected by Notification type. Weak evidence suggests that Notification type may have had an influence on ratings associated with General SA ( $F(2,21) = 2.301$ ,  $p = 0.074$ ). Tukey post-hoc tests show the only pairwise comparison nearing significance ( $p = 0.064$ ) to be that indicating improved General SA when crews had AAA Notifications than VVV Notifications.

Binomial tests with the mid-scale value as the cut point for equal likelihood of responses tested permits a more sensitive assessment for each Notification condition. Responses to the Workload questions were not biased to either side of the scale for AAA or VVV conditions, but shows a weak bias to the end of the scale suggesting an increase in workload for the VAV condition ( $p = 0.07$ ). There seemed to be increased awareness of arrival speeds in the AAA condition ( $p = 0.008$ ), moderate evidence for this same trend for the VAV condition ( $p = 0.07$ ), and no significant evidence of a bias in scale use for the VVV condition ( $p = 0.289$ ). With regard to general situation awareness, no bias in scale responses is evident for the VAV ( $p = 0.289$ ) or VVV ( $p = 1.00$ ), but a weak result indicates greater general situation

awareness in the AAA condition ( $p = 0.07$ ) - as reflected in the aforementioned parametric test. Responses for all conditions demonstrate indications that current crew coordination and cockpit resource management strategies are appropriate for FIM operations; these are strong indications for the AAA and VVV conditions (both  $p = 0.008$ ), and weaker for the VAV condition ( $p = 0.07$ ).

### Comments on Operational Impact

Subjects' comments regarding workload, situation awareness, and crew coordination were generally positive. Where concerns were mentioned, these centered on increased distraction from outside scan, the potential for obtaining speeds that are difficult to achieve, and the impact on passenger comfort and fuel use should speed brakes need to be used more aggressively than in current operations. While subjects generally reported that standard crew coordination and resource management techniques serve them well in conducting FIM operations, several did mention the need to emphasize callouts when a new speed target is provided and when speed deviations are detected.

## Discussion

### Avionics Conditions

While the primary flight display formatting in the integrated condition, and the placement of the Fire-EFB condition could be improved, both these conditions were generally acceptable to these subjects, as indicated by both commentary and paired comparison data. Ratings of Operational Acceptability when split by Notification type showed insignificant differences among these Avionics conditions, but the differences observed when subjects received aural indications were consistent with commentary and paired comparison data. Several subjects mentioned that they thought that they would prefer the condition with the AGD, but found that the indications on it lacked salience, and so reliance on it resulted in missing speed targets. This suggests that focal distance, in addition to eccentricity be considered when assessing useful "field of view." The condition, in which the EFB was mounted off to the side, was clearly unacceptable to most subjects. In this study, charts were also



implemented on the EFB, and so there was some commentary regarding the shared use of a device to present both FIM information and charts. Data related to this topic will be reviewed in a forthcoming paper.

### ***Notification Methods***

The particular aural tone used in this study seemed to be found appropriate by subjects. It received middling ratings with regard to perceived urgency, and lower than middling responses when rated on how annoying its tonal qualities were. It appeared to be salient enough to capture attention in most cases, but not so demanding that it was impossible for other flight related tasks to preclude noticing it on first pass. Subsequent reminders were generally noticed in a timely manner if the first indication was missed.

Aural indications were preferred by roughly 92% subjects in this study for at least one (and for half of the subjects, all) of the FIM events. Even the majority of those subjects who received only visual indications suggested that aural indications would be appropriate for FIM events.

This study used the same aural indication for all the FIM events. The concept was that the tone would cue the crew to direct their attention to the FIM operation and displays, without differentiating among events. Several subjects indicated that either the tone should only be used for a subset of these events, or that different tones, or voiced messages should be used to cue the type of FIM event. The preference for aural indications by these subjects and the recommendation by a few to multiply this for each type of event runs counter to concerns regarding proliferation of aural indications on modern commercial flight decks [3]. Subjective data, while valuable, has been known to dissociate from performance data [4]. As such, future analyses from this study's data will assess crew flight path deviation performance (speed and vertical deviations) and responsiveness to FIM events.

Subjects who received aural indications for all three of the FIM events provided significantly higher ratings than the other two groups (who did not experience them) for the need for aural indications for conformance deviations. Results also indicate a

weak effect suggesting that subjects who received aural indications for all FIM events felt increased situation awareness in other aspects of the flight. There is also some indication that those who received aural indications for conformance deviations felt they had improved situation awareness of arrival speeds over their typical operations. While it is tempting to be congratulatory regarding the increase in situation awareness of arrival speeds, it is always important to wonder "at what cost?" This is encouraging, but additional study with more objective and sensitive measurements addressing situation awareness content is required to validate this tenuous finding. Oculometer data taken during this study may further inform the impact of these conditions on scan, and therefore at least the constituent information acquisition for building situation awareness.

### ***Reminder & Conformance Thresholds***

Subjective reports associated with the logic for reminders and conformance indications showed consistency with the logic implemented in this study. Subjects were more likely to feel that the conformance indications were false alarms than were reminder indications. In observation, we note that there were in fact false alarms; that is, occasions in which the subjects did respond in a timely manner and still received an indication that they were deviating from the profile speed in excess of expectations. The deleterious effects of false alerts, in terms of reducing compliance, are well-known [5], and should motivate more sophisticated consideration of thresholds. We suggest that different speed deviation thresholds may be appropriate for different regions of flight, as would be consistent with pilots' expectations.

### ***Operational Acceptability***

Subjects were generally very enthusiastic about the concept of IM. Ratings indicated only improvements in general situation awareness (for those receiving aural indications for all events), and for arrival speeds (for those receiving aural for all and for conformance deviation events). Ratings also suggest no significant impact regarding crew coordination strategies or motivation for significant training to support IM. Workload was rated to be comparable to non-IM operations. A slightly troubling, and perplexing finding is a slight tendency

for subjects who received aural indications only for conformance deviations to feel increased workload.

Taken with the result that subjects experiencing the VAV Notification method reported more urgency of the tone than those who experienced the tone for all three FIM events (the AAA method), we suggest that perhaps the fact that AAA subjects experienced more tones, they had a stronger and less startling conditioned response to associate the tone with attending to FIM equipment. Another explanation may derive from the framing of the tone's meaning: in the AAA method, the tone usually conveyed the presence of a new speed target – and was therefore non-pejorative; whereas in the VAV condition, the tone was always associated with having made an error in adherence to the speed profile. Some subjects' requests for differentiated tones for the different IM events, and poses an interesting question for further experimentation. To the degree that aural indications convey the information required for action, they obviate the need to reference visual information; but to the degree that they must still be interpreted, impart more reliance on memory to decode. When tones are mode awareness indications, they are less specific and require reference to visual indications, but they are more frequent and reinforced referents to the appropriate equipage. In this case, while presenting fewer tones to the flightdeck environment, there is a possibility of mistaking a tone that is issued to announce a new commanded speed for one that is subsequent to a known speed profile conformance deviation.

### ***Caveats***

The reader is to be reminded that the results obtained are inextricably linked to the particular implementations under test. For example, we found that certain visual indications lacked in 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 manipulation is a between-subject variable, and so received fewer data points per subject than the within-subject variable testing Avionics conditions. This paper focuses on the summary results obtained by Post-Experiment questionnaires. These limitations will be mitigated by consideration of performance data and

Post-Run questionnaire data which will provide more sensitive measurement of these effects.

## **References**

- [1] 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.
- [2] 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 102 (2): 251–318. Retrieved 2013-6-22.
- [3] Boucek, G. P., D. C. Hanson, B.L. Berson, M. F. Leffler, D. A. Po-Chedley, and J. F. Hendrickson, 1981, "Aircraft Alerting Systems Standardization Study," AIAA/IEEE Digital Avionics Systems Conference, AIAA: New York, 119-128.
- [4] Andre, A.D. and Wickens, C.D. (1995), "When Users Want What's Not Best for Them." Ergonomics in Design, October, 10-14.
- [5] Geels-Blair, K., Rice, S., and 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.

## **Acknowledgements**

This work was conducted under collaborative sponsorship by the NASA Airspace Systems Program and the Federal Aviation Administration, Human Factors Division (ANG-C1).

## **Email Addresses**

K.A.Latorella@nasa.gov

*32nd Digital Avionics Systems Conference*

*October 6-10, 2013*