Usability Evaluation of Indicators of Energy-Related Problems in Commercial Airline Flight Decks

Emory T. Evans
Flight Software Systems Branch
NASA Langley Research Center
Hampton, Virginia
e.t.evans@nasa.gov

Lynda J. Kramer
Crew Systems and Aviation
Operations Branch
NASA Langley Research Center
Hampton, Virginia
lynda.j.kramer@nasa.gov

Timothy J. Etherington
Technical Fellow
Collins Aerospace
Hampton, Virginia
timothy.j.etherington@nasa.gov

Taumi S. Daniels
Electronic Instrument Systems
Branch
NASA Langley Research Center
Hampton, Virginia
taumi.s.daniels@nasa.gov

Steven D. Young
Safety-Critical Avionics Systems Branch
NASA Langley Research Center
Hampton, Virginia
steven.d.young@nasa.gov

Yamira Santiago-Espada
Safety-Critical Avionics Systems Branch
NASA Langley Research Center
Hampton, Virginia
yamirah.santiago-espada@nasa.gov

James R. Barnes
Simulation Development & Analysis
Branch
Hampton, Virginia
james.r.barnes @nasa.gov

Abstract— **A series of pilot-in-the-loop flight simulation studies were conducted at NASA Langley Research Center to evaluate indicators aimed at supporting the flight crew's awareness of problems related to energy states. Indicators were evaluated utilizing state-of-the-art flight deck systems such as on commercial air transport aircraft. This paper presents results for four technologies: (1) conventional primary flight display speed cues, (2) an enhanced airspeed control indicator, (3) a synthetic vision baseline that provides a flight path vector, speed error, and an acceleration cue, and (4) an aural airspeed alert that triggers when current airspeed deviates beyond a specified threshold from the selected airspeed.**

Full-mission high-fidelity flight simulation studies were conducted using commercial airline crews. Crews were paired by airline for common crew resource management procedures and protocols. Scenarios spanned a range of complex conditions while emulating several causal factors reported in recent accidents involving loss of energy state awareness by pilots.

Data collection included questionnaires administered at the completion of flight scenarios, aircraft state data, audio/video recordings of flight crew, eye tracking, pilot control inputs, and researcher observations. Questionnaire response data included subjective measures of workload, situation awareness, complexity, usability, and acceptability. This paper reports relevant findings derived from subjective measures as well as quantitative measures.

Keywords—Energy state, flight deck automation, aircraft state awareness, decision making, aviation safety, flight deck simulation

I. INTRODUCTION

In a series of planned experiments (referred to as the Automation and Information Management Experiments, or AIME), technology concepts to promote aircraft state awareness (ASA) are being evaluated at NASA Langley Research Center. A Boeing 787 (B-787) flight deck configuration was chosen as the reference platform onto which the ASA technologies would be inserted. Data collection for AIME 1 (late 2015 – early 2016) [1 - 6], AIME 2 (mid 2018) [7 - 11], and AIME 2.5 (late 2018 – early 2019) have been completed. Data collection for AIME 3 is planned for summer/fall 2019. The goal of these experiments is to demonstrate and evaluate technologies to improve ASA in a relevant environment across a span of conditions associated with potential loss of ASA.

Over the course of the AIME experiments completed to date, the following technologies have been evaluated:

- Trajectory Prediction System (TPS) on the Navigation Display (ND) and Vertical Situation Display (VSD) *AIME 1, 2.5*

- Predictive Alerting of Energy (PAE) problems on Engine-Indicating and Crew Alerting System (EICAS), ND, and VSD. *AIME 1, 2.5*

- System Interaction Synoptic (SIS)/enhanced Synoptics (eSyn) on Synoptics display with simplified electronic checklists (sECL). *AIME 1, 2, 2.5*

- Maneuver Envelope (ME) estimation on the Primary Flight Display (PFD). *AIME 1*

- Enhanced Airspeed Control Indicator (EACI); on the PFD. *AIME 2*

- Stall Recovery Guidance (SRG) on the PFD. *AIME 2*

- Flight path vector (FPV), speed error tape, and acceleration caret (HUD-type symbology) on a Synthetic Vision System (SVS) PFD [referred to as "FPV" configuration]. *AIME 2.5*

- Aural Airspeed Alerting (AAA) in conjunction with FPV configuration *AIME 2.5*

This paper compares results for EACI, FPV, and AAA ASA technology concepts to a conventional PFD (baseline condition or "BL" in Fig. 1) across similar experimental scenarios with loss of airspeed events.

II. TECHNOLOGIES DISCUSSED

The "EACI" configuration is an enhancement to the conventional PFD airspeed tape (Fig. 2) and indicates whether the auto-flight system (including auto-throttles) is actively controlling to a target airspeed [12]. EACI was implemented as additional display elements on a standard blue-over-brown PFD. When airspeed is not actively controlled by the automation, whites XX's are shown:

- at the commanded speed indicator at the top of the airspeed tape,

- at the boxed current airspeed indicator in the center of the airspeed tape, and

- at the airspeed bug location as set in the MCP or FMS.

The FPV configuration augmented a conventional PFD with HUD-type symbology (flight path vector, speed error tape, and acceleration caret) and synthetic vision imagery (Fig. 3). This SVS PFD with HUD-type symbology concept has been studied extensively as a mitigation for spatial disorientation incidents/accident [13 – 15], but it has not been specifically tested for mitigating loss of airspeed awareness.

The "AAA" configuration triggers an "airspeed" or "airspeed, airspeed" aural callout when current airspeed deviates from the selected airspeed by 10 knots or more than 20 knots, respectively. AAA was implemented in conjunction with the FPV configuration (flight path vector, speed error tape, acceleration caret on an SVS PFD).

III. OVERVIEW

A. Facilities

The AIME experiments were conducted in the Research Flight Deck (RFD) within the Cockpit Motion Facility (CMF) at NASA Langley Research Center (Fig. 4). The RFD is a high fidelity, full-motion flight simulator and was configured to emulate B-787 displays.

B. Flight Deck Layout

Boeing 787-like flight deck displays and layout served as the reference condition (Fig. 5). The Lower Multi-Function Display (LMFD), served primarily as the pilot interface to the Flight Management System (FMS). Air Traffic Control (ATC) data link messages were displayed inboard of the Electronic Flight Bags (EFB). EACI, FPV, and AAA could be activated to supplement this baseline.

Fig. 1. BL display at 10 kt airspeed loss after autothrottle failure.

Fig. 2. EACI display at 10 kt airspeed loss after autothrottle failure. White XX's indicate automation is not controlling to target airspeed.

Fig. 3. FPV display at 10 kt airspeed loss after autothrottle failure. Note acceleration caret and speed deviation box below left wing of FPV symbol.

Fig. 4. RFD interior.

Fig. 5. AIME flight deck display layout.

C. Test Subjects

For the AIME experiments, commercial airline pilots flew scripted scenarios in the RFD. Captains and First Officers from the same airline were paired to ensure common operational procedures. The Captain flew in the left seat of the flight deck and the First Officer flew in the right for the duration of the test. Each crew began with a few runs for familiarization with the simulator and the new technologies. Data collection runs followed these training runs [16]. Data were collected individually from the pilot flying (PF) and the pilot monitoring (PM). The Captain and First Officer alternated PF/PM roles periodically so that each pilot assumed each role for about half the scenarios. Roles were not changed during flights, only between.

D. Scenarios

AIME flight scenarios spanned a range of conditions designed to help expose state awareness issues where the technologies under evaluation could prove useful. The scenarios were designed to emulate some of the causal factors reported in accidents involving loss of ASA. They were also intended to immerse flight crews in a complex operational environment with high-density traffic, adverse weather, both voice and data link ATC communications, RNAV and RNP arrivals, and off-nominal events. The off-nominals were intended to help expose state awareness issues and/or to mimic relevant accidents/incidents involving loss of ASA. Scenario flight times were about 15-20 minutes. There were three AIME 2 and three AIME 2.5 scenarios designed to induce low energy events. The salient events in these scenarios are noted below:

AIME 2 Scenario 3a. Autothrottle (A/T) fails on approach
- Inbound to JFK ILS 4R
- Cleared for approach
- At 1200 ft Above Ground Level (AGL), A/T fails silently
- Run ends at go around or landing

AIME 2 Scenario 3b. A/T fails in hold
- Inbound to JFK RNAV 13L
- Message to hold at CAMRN
- Descending through 11500 ft Mean Sea Level (MSL) to enter hold, A/T fails silently
- Concurrent distraction message for runway change to ILS 13L
- Run ends when crew detects A/T not engaged

AIME 2 Scenario 4b. A/T fails in cruise
- Takeoff from DCA
- Inbound to JFK RNAV 13L
- In cruise, A/T fails silently
- Concurrent distraction message for traffic ahead
- Run ends when crew notices A/T not engaged

AIME 2.5 Scenario 5a. A/T retards due to radio altimeter (RA) failure on approach
- Inbound to JFK ILS 4R
- At 1100 ft AGL, RA reports -8 ft.
- Rollout flare modes armed, so A/T retards
- Concurrent distraction message for rejected takeoff ahead on 13L
- Run ends at go around or landing

AIME 2.5 Scenario 7a. A/T fail (off-nominal event 1) and blocked pitot system (off-nominal event 2) in hold

- Inbound to JFK RNAV 13L
- Message to hold at CAMRN
- Descending through 11500 ft MSL to enter hold, A/T fails silently (event 1)
- Concurrent distraction message for runway change to ILS 4R
- Once crew detects A/T not engaged (end of event 1), principal investigator triggers a datalink message (initiation of off-nominal event 2) to descend in the hold
- At 9000 ft MSL, total pitot system blockage (event 2)
- EICAS message for unreliable airspeed
- AOA speed indicated on Captain's PFD
- Inoperative: flight directors, autopilot, A/T
- Run ends when crew completes unreliable airspeed checklist

AIME 2.5 Scenario 7b. A/T fails in cruise
- Takeoff from DCA
- Inbound to JFK RNAV 13L
- In cruise, A/T fails silently
- Concurrent distraction message for traffic ahead
- Run ends when crew recognizes A/T not engaged

AIME 2 scenario 4b and AIME 2.5 scenario 7b are the same scenario. AIME 2 scenario 3a and AIME 2.5 scenario 5a are designed to emulate the same accident. In 3a the autothrottle silently fails with little pitch movement. It was desired to test the symbology without using significant display movement. The limited pitch movement results in a slower speed bleed off and more time to check the error. In 5a, the radio altitude failure triggers a flight control mode change and the autoland systems transitions into flare and idle, even though the aircraft is still airborne. This is the actual accident scenario. Pitch movement is noticeable but airspeed bleeds off much more quickly resulting in less time to catch the error before stall warning. AIME 2 scenario 3b and AIME 2.5 scenario 7a are exactly the same through autothrottle failure and recovery. For 7a an additional pitot failure was inserted at the end of the scenario to test synoptic and checklist functions. This allows for a between subjects comparison of airspeed awareness results across AIME 2 and 2.5 experiments with the exception of some complexity and workload measurements due to the additional failures.

E. Data Collection

During each flight, objective data were recorded including aircraft state parameters, audio, video, and eye tracking. In addition, for the scenarios described in Section III D, the researcher noted the elapsed time from A/T failure until pilot recognition of this event (i.e., began manually controlling A/T). After each flight, subjective data were collected from each pilot including perceived complexity, workload, situational awareness, and the usability and acceptability of the technology concepts. After the last flight for each crew, pilot comments and feedback were gathered in a summary debrief. In the sections below, relevant findings from the objective and subjective data are reported. Representative pilot comments on the technologies are provided as well.

IV. RESULTS AND DISCUSION

The scenarios used to evaluate EACI, FPV, and AAA were flown both with and without the technology active. Table I below shows the distribution of completed questionnaires (Q's) for the four technology configurations.

TABLE I. COMPLETED POST-RUN QUESTIONNAIRES

AIME	Scenario	Tech	Q's PF+PM
AIME 2	3a - Approach	EACI	18
		BL	8
	3b - Hold	EACI	14
		BL	12
	4b - Cruise	EACI	20
		BL	6
AIME 2.5	5a - Approach	AAA	14
		FPV	6
	7a - Hold	AAA	14
		FPV	6
	7b -Cruise	AAA	14
		FPV	4

A. Pilot recognition of airpeed loss

The following discussion assumes the addition of synthetic vision to the display format from AIME 2 to AIME 2.5 does not contribute in a meaningful way during the scenarios tested and that the better airspeed awareness was due to symbology and aural cues alone. There are two components that may have an interaction. The scenarios utilized thrust loss with an associated reduction in speed. Research shows that textured synthetic vision displays can be effective in determining ego speed but the effect is minimal at higher altitudes such as those used in the testing. Pitch changes may be more noticeable with SVS, and the pitch attitude did increase with airspeed loss while maintaining altitude. This pitch change was slow and subtle and not felt to contribute to awareness of airspeed loss.

The median airspeed loss for the four configurations tested during the hold and cruise scenarios is shown in Fig. 4. The median airspeed loss for EACI, FPV, and AAA were lower than the BL condition (EACI = 16 kt, FPV = 13 kt, AAA = 13 kt, BL =25 kt). Moreover, the spread across the tests may reflect the pros and cons of each technology concept. For example, the spread when using EACI matches that of BL. This is likely because it is so close to the BL design and a change to scanning behavior is not required. However, to be effective the X's must be noticed and processed and that is where visual-only cues can be problematic. In two cases when using EACI, pilots did not recognize the speed loss until stick shaker was activated [8]. These data points are shown outside the BL-like spread. Subjective comments indicate that the EACI white X's might be more noticeable by using an additional alerting method (e.g., flashing the X's, aural alerts). On the positive side, since EACI showed the loss of autothrottle immediately before the aircraft even started to decelerate, some pilots noticed the mode change and immediately corrected the problem.
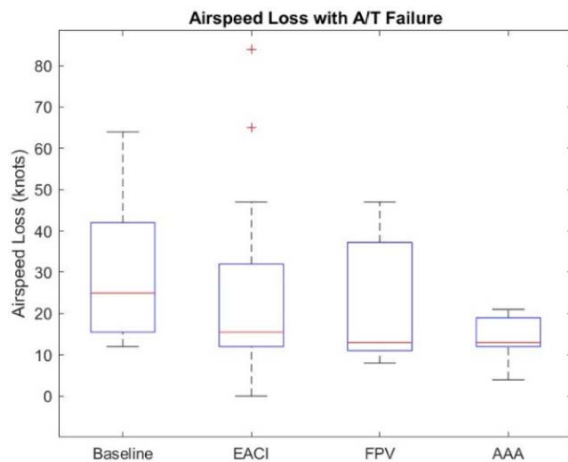

Fig. 4. Median airspeed loss collapsed across hold and cruise scenarios 3b, 4b, 7a, and 7b.

The spread when using the HUD-like symbology (referred to as FPV) suggests quicker pilot recognition of airspeed loss, but some of its potential benefit is likely hidden due to limited training and/or experience by the crews with HUD-like symbology such as a flight path vector. As a result, this is likely a worst-case spread that would shrink with training and experience with its use. Since the FPV shows energy loss due to deceleration before a perceptible difference between current speed and selected speed, it may result in quicker detection than audio-based alerting triggered when current airspeed deviates beyond a specified threshold (e.g., 10 kt) from the selected airspeed. Even though the HUD-like symbology is visual, it is an effective mitigation due to the nature of the location of the symbology. That is the primary indication of the flight of the aircraft and when monitoring the automation, the FPV is in the area with the most critical information of how well the automation is tracking the desired aircraft parameters. It is in the primary field of vision, even when monitoring the automation.

The smallest spread and median value was observed using AAA. In contrast to FPV, this may represent a best-case spread as very limited training is required to understand and utilize the concept. The error value was carefully designed to eliminate false alarms, even in turbulence, when the automation is controlling speed. There were few false alarms noted during testing, even during hand flown segments such as the departure, but further study may be required. These cons associated with audible alerts could increase the spread and/or possibility of missed recognitions but the audio component here was also persistent after every 10 knot deviation. This design is intended to reduce the probability of missed detection leading to a catastrophic event since multiple audio alerts are provided for the event. One negative aspect, but maybe not significantly relevant operationally, is the audio alert is designed with a threshold to prevent nuisance alerts. Even in some turbulence cases the 10 knot threshold would be exceeded with the auto throttle engaged. Detection below 10 knots is likely due solely to mode recognition. Due to aircraft dynamics, one test run resulted in the first aural airspeed alert to be triggered at 20 knots deviation from commanded airspeed instead of the 10 knot deviation trigger used in the other runs.

B. Complexity, Workload, and Situational Awareness

1) Complexity

After each flight, pilots rated the complexity of four elements: (1) the task, (2) operational environment, (3) system/automation, (4) information provided. Each factor was rated from 1 (not complex) to 10 (extremely complex). Because the rating for each element tended to track with the other elements for a given scenario, the four elements are averaged below into a single rating for simplicity.

Table II shows the average complexity ratings for the four technology configurations.

TABLE II. AVERAGE COMPLEXITY

Scenario	BL	EACI	FPV	AAA
App (3a/5a)	5.4	5.2	5.7	6.3
Hold (3b/7a)	4.9	5.3	6.0	7.1
Cruise (4b/7b)	4.9	3.6	3.5	3.0

Average complexity ranged from a low of 3.0 for scenario 7b to a high of 7.1 for scenario 7a, which

exposed the pilots to cascaded system failures. Although scenarios 3b and 7a (through event 1) are largely equivalent, in 7a crews had the additional burden of having to deal with unreliable airspeed information due to blocked pitot tubes.

As described in Section III D, one of the research objectives was to test the technologies in a complex environment. These scores validate that this research objective was met.

2) Workload (WL)

Perceived workload was measured using the NASA Raw Task Load Index (RTLX) [17 - 19].

RTLX derives an overall workload score based on the unweighted average of ratings of six subscales (mental demand, physical demand, temporal demand, performance, effort, and frustration level). Each subscale rating, except performance, is scored on a scale of 0 (very low) to 100 (very high). Performance is scored on a scale of 0 (perfect) to 100 (failure).

Tables III shows the mean perceived workload for the AIME 2 and 2.5 low-energy scenarios.

TABLE II. WORKLOAD

Scenario	BL	EACI	FPV	AAA
App (3a/5a)	43.9	40.9	49.7	57.7
Hold (3b/7a)	48.1	45.8	50.7	58.9
Cruise (4b/7b)	38.8	34.5	13.9	26.2

Workload ranged from 13.9 (low workload) for scenario 7b to 58.9 (moderate workload) for scenario 7a, with its cascaded system failures. Moderate workload levels (value between 40 and 60) were observed for all 4 technologies in the approach and hold scenarios. Low (value less than 20) to moderately low (value between 20 and 40) workload levels were observed for the technologies during the cruise scenario.

3) Situational Awareness

Situational awareness (SA) was measured by using the 3-D Situational Awareness Rating Technique (SART) [20]. It calculates SA based on three dimensions: (1) Demand on pilot attentional resources, (2) Supply of pilot attentional resources, and (3) Understanding by the pilot of the situation. Pilots rated these after each flight on a scale of 0 (very low) to 100 (very high). From these, SA is calculated as:

$$SA = Understanding - (Demand - Supply)$$

Table IV shows SA for the AIME 2 and 2.5 low-energy scenarios.

TABLE IV. SITUATIONAL AWARENESS

Scenario	BL	EACI	FPV	AAA
App (3a/5a)	75.5	69.9	61.8	21.9
Hold (3b/7a)	59.0	81.1	60.5	9.4
Cruise (4b/7b)	74.3	96.6	130.3	110.7

Situational awareness ranged from a high of 130.3 with the FPV technology in cruise scenario 7b to a low of 9.4 for AAA in hold scenario 7a, with its cascaded system failures.

Augmenting the PFD with EACI symbology led to increased SA compared to the baseline condition for the hold and cruise scenarios. The approach scenario with A/T failure shows an apparent decrease in SA with EACI compared to baseline condition but these differences are not operationally significant.

Similarly, augmenting the conventional PFD, with HUD-like symbology (FPV configuration), led to increased SA compared to baseline for the hold and cruise scenarios. HUD-like symbology did not appear to provide any improvements with SA over the baseline for the approach scenario.

SA was lower with AAA than FPV in all scenarios. This result may have been due to the pilots reacting to aural airspeed alert by immediately manually controlling speed with A/T before determining what led to the decay in airspeed.

Finally, the scenario 7a AAA results for complexity, workload, and situational awareness were noticeably affected by combining two negative events (A/T failure and unreliable airspeed) in one run. This is a lesson learned for the research team.

C. Usability and Acceptability

After each EACI or AAA flight, the PF and the PM independently rated both the acceptability and the usability of the technology used in that scenario. Usability and acceptability were not assessed for BL and FPV.

After every scenario flown, each pilot rated the acceptability of the subject technology on a 1 to 7 scale. Where

1 = Very unacceptable. I did not like the technology and would not use it in normal operations.

4 = Average. I liked the technology and would use it in normal operations, but would like to see some improvements.

7 = Very acceptable. I like the technology very much and would use it without any improvements.

Usability was assessed with the System Usability Scale (SUS) [21 - 23]. SUS scores can range from 0 to 100, but they are not percentile ranks. However, they can be converted to percentile ranks. As a way to intuitively interpret SUS scores, [24] proposed a correlation between SUS scores and letter grades. Based on extensive research [23] showed specific letter grades can be associated with corresponding SUS score ranges. This mapping is used here to put the SUS scores in the context of letter grades.

Usability and acceptability results for common scenarios of A/T failure in hold and cruise are shown in Fig. 5. SUS scores, means with ± 2 standard errors (SE), are shown as multi-color bar graphs (referencing left y-axis scale) and acceptability scores are shown as red circles with mean values listed adjacent to them (referencing right y-axis scale).Self-reported post-run ratings indicated that pilots judged the pilot interfaces to be usable and acceptable for the EACI and AAA concepts for hold and cruise operations.
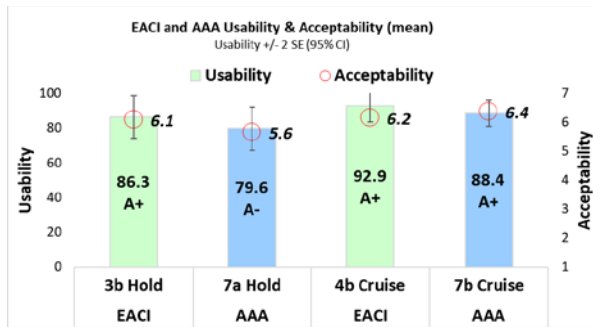

Fig. 5. Comparing EACI and AAA across common scenarios

Across all three scenarios (approach, hold, cruise), EACI usability was rated A+ by both PF and PM. EACI acceptability was rated high, with a 6.2 by PF and 6.0 by PM. AAA was rated more useful and acceptable for the PM than the PF. Usability was A+ for PM but B for PF. Acceptability was rated 6.4 by PM but 5.3 by PF. More training and/or work may be needed to mature the AAA technology to the benefit of the PF and the scores may be confounded with the additional failures encountered with the AAA scenarios.

D. Pilot comments

Discussions between pilots, and between pilots and researchers, were held after each run and during the post-experiment debriefs. Representative comments for EACI and AAA are noted below.

EACI

1. …very intuitive to see. Maybe add an aural warning or tone.

2. …good idea but not extremely noticeable; needs to be more "in your face"

3. …Good - maybe make the "X" amber or red

4. …make the XX's rapid flashing

5. …would like something less subtle such as "A/T OFF" in center of the PFD

6. … very useful; good to get a visual picture that you have a system failure.

7. …make the trend airspeed arrow change color

8. …really liked it; clear, and alerted me quickly to uncontrolled airspeed.

9. …very useful for understanding the autothrottle state

10. …the white Xs stood out and made it much easier and quicker to recognize the auto-throttles weren't working.

11. …as quick or quicker noticeability compared to FMA; very user friendly.

12. …great visual indication for failed autothrottles; add XXs to the FMA as well.

13. …great safety improvement; but add an EICAS message.

AAA

1. …very effective in the "heat of battle;" provides a extra level of safety.

2. …good for catching failures where thrust is in hold mode or fails…essential for hand flying

3. …outstanding- excellent means to alert crew to low or high energy state; will reduce low-speed events with A/T disconnected or airspeed inop or set incorrectly

4. ...simple but extremely helpful; you instantly focus on airspeed; it would be nice to have the FPV/ISEC on the PFD; overall a great system.

5. …the escalation of alerts provided a sense of urgency. The tone, volume and voice got my attention to focus on the FPV and speed trend arrow.

6. …I'm a believer in this but the alerts need a recognizable hierarchy of alerts and commands, e.g., "airspeed" followed by "airspeed low".

7. …invaluable.

8. …definitely got my attention; suggest the volume go higher at each level; voice was too smooth and calm.

V. SUMMARY OF RESULTS

Results for EACI, FPV, and AAA compared to BL were presented in three areas:

1. Mean airspeed loss prior to pilot recognition was lower for all three ASA technologies compared to BL. With our limited number of samples, the ASA technologies on a PFD with HUD-type symbology appear to be more effective for airspeed awareness than conventional PFD without HUD-type symbology

   - BL airspeed loss was 25 kt.
   - EACI airspeed loss was 16 kt.
   - FPV airspeed loss was 13 kt.
   - AAA airspeed loss was 13 kt.

   By this measure, the AAA and FPV ASA concepts employing HUD-type symbology outperformed EACI, which did not employ it. In two cases using EACI pilots did not recognize airspeed loss until stick shaker activation. Current PFDs have a lot of information on them and visual-only enhancements may not be completely effective for energy state awareness. The BL configuration noticeably underperformed the ASA technologies for this measure.

2. Complexity, workload, and situational awareness as perceived by the pilot were assessed for EACI, FPV, AAA, and BL.
   - Complexity scores appear to validate that the high-fidelity full-mission flight simulation was able to replicate typical crew procedures and workload as well as complex off-nominal situations as encountered during accidents.
   - Moderate workload levels were reported across each of the technologies
   - EACI and FPV improved SA for the pilots over the baseline condition.
   - It appears that although pilots recognized airspeed loss most quickly with AAA that it had a negative effect on their SA compared to baseline. They quickly remedied the loss of airspeed when the aural alert was activated but may not have had an awareness of what caused the airspeed loss or they may have rated a lack of SA because AAA did not help in the failure encountered. No stick shaker events were encountered with the AAA or FPV technology.

3. Usability and acceptability were assessed for EACI and AAA
   - EACI was rated highly by both PF and PM
     - Usability was rated A+ by both
     - Acceptability was rated 6.2 by PF and 6.0 by PM.
   - AAA was rated higher by the PM than the PF.
     - Usability was rated A+ by PM but rated B by PF.
     - Acceptability was rated 6.4 by PM but rated 5.3 by PF.

Both EACI and AAA were rated usable and acceptable by the pilots.

VI. CONCLUDING REMARKS

All the technologies tested outperformed the baseline and provided mitigations for detecting issues with energy awareness.

FPV performed well in this study but the results need to be tempered due to the small number of runs in the current study. The results for airspeed loss appear to validate the Commercial Aviation Safety Team (CAST) recommendation that HUD-like symbology may be one of the most effective mitigations to loss of energy awareness. Where possible, flight path based symbology should be strongly considered in light deck designs instead of the standard attitude-based displays in common use today.

Aural alerts have been used effectively to minimize altitude clearance violations and appear to be a simple and effective solution to airspeed awareness. The implementation used in the study was effective because it was persistent, unlike other aural and visual warning that present only once. It helps flight crews monitor the automation and supports the current visual symbology.

EACI is effective when noticed. The additional visual elements and changing symbology were very effective if pilots noticed the changes when they occurred. There were times were the visual elements just blended into the symbology already present and the airspeed loss was not detected until additional warnings like stick shaker were activated.

REFERENCES

[1] E. Evans, S. Young, T. Daniels, Y. Santiago-Espada, "Analysis of Pilot Feedback Regarding the Use of State Awareness Technologies During Complex Situations," Proceedings of the 35th AIAA/IEEE Digital Avionics Systems Conference, Sacramento, CA, Sept 2016.

[2] S. Young, T. Daniels, E. Evans, K. Shish, S. Schuet, T. Etherington, M. Uijt de Haag, and D. Kiggins, "Evaluating Technologies for Improved Airplane State Awareness and Prediction", AIAA Science and Technology Forum and Exposition (SciTech), San Diego, CA, January 4-8, 2016.

[3] S. Young, T. Daniels, E. Evans, E. Dill, M. Uijt de Haag, T. Etherington, "Flight Simulation Study of Airplane State

Awareness and Prediction Technologies," Proceedings of the 35th AIAA/IEEE Digital Avionics Systems Conference, Sacramento, CA, Sept 2016.

[4] M. Uijt de Haag et. al.; "Energy State Prediction Methods for Airplane State Awareness," Proceedings of the AIAA/IEEE Digital Avionics Systems Conference, Sacramento, CA, September 25-29, 2016.

[5] T. Lombaerts, S. Schuet, D. Acosta, J. Kaneshige, and K. Shish, "Piloted Simulator Evaluation of Maneuvering Envelope Information for Flight Crew Awareness," paper no. AIAA 2015-1546, AIAA SciTech 2015, Kissimmee, FL, Jan 2015.

[6] K. Shish, J. Kaneshige, D. Acosta, S. Schuet, T. Lombaerts, L. Martin, and A. Madavan, "Trajectory Prediction and Alerting for Aircraft Mode and Energy State Awareness," paper no. AIAA 2015-1113, AIAA SciTech 2015, Kissimmee, FL, Jan 2015.

[7] Young, S. D., "Airplane State Awareness and Prediction Technologies – Research Overview (2014-2018)," in AIAA Science and Technology Forum and Exposition SciTech, 2019.

[8] Kramer, L. J., Etherington, T. J., Evans, E. T., Daniels, T. S., Young, S. D., and Barnes, J. R., "Evaluation of Technology Concepts for Energy, Automation, and System State Awareness in Commercial Airline Flight Decks," in AIAA Science and Technology Forum and Exposition SciTech, 2019.

[9] Daniels, T. S., Ferguson, C. M., Dangtran, E. T., Korovin, R. M., Kramer, L. J., Evans, E. T., Santiago-Espada, Y., Kiggins, D., Etherington, T. J., and Barnes, J. R., "Regarding Pilot Usage of Display Technologies for Improving Awareness of Aircraft System States," in IEEE Aerospace Conference, Yellowstone Conference Center, Big Sky, Montana, 2019.

[10] Schuet, S., Lombaerts, T., Stepanyan, V., Kaneshige, J., Hardy, G., Shish, K., Robinson, P., Etherington, T., Kramer, L. J., Evans, E. T., Daniels, T. S., and Young, S. D., "Piloted Simulation Study Findings on Stall Recovery Guidance," in AIAA Science and Technology Forum and Exposition SciTech, 2019.

[11] Lombaerts, T., Schuet, S., Stepanyan, V., Kaneshige, J., Hardy, G., Shish, K., Robinson, P., Kramer, L., Etherington, T., Daniels, T., Evans, E. T.,and Young, S., "Design and Piloted Simulator Evaluation Results of Model Independent Stall Recovery Guidance," in AIAA Science and Technology Forum and Exposition SciTech, 2019.

[12] Sherry, L. and Mauro, R., "Design of Cockpit Displays to Explicitly Support Flight Crew Intervention Tasks," in Digital Avionics Systems Conference (DASC), 2014 IEEE/AIAA 33rd, 2014, pp. 2B5.1–2B5.13

[13] K.K.E. Ellis, L.J. Prinzel, J.J. Arthur, S.M. Nicholas, D. Kiggins, H. Verstynen, C. Hubbs, J. Wilkerson, "Design considerations for attitude state awareness and prevention of entry into unusual attitudes", 19th International Symposium on Aviation Psychology (2017), 8–11 May 2017, Dayton, OH

[14] Prinzel, L.J., Arthur, J.J, Nicholas, S.N., Lake, R.C., Ballard, K., and Kiggins, D, "Evaluating synthetic vision displays for enhanced airplane state awareness," Proc. SPIE10642, 10642D (2018).

[15] Prinzel, L.J., Ellis, K.K.E.,Arthur, J.J, Nicholas, S.N., and Kiggins, D, "Synthetic Vision System Commercial Aircraft Flight Deck Display Technologies For Unusual Attitude Recovery", in International Symposium on Aviation Psychology, Dayton, OH, 2017.

[16] T. Etherington, L. J. Kramer, K. Shish, S. D. Young, E. T. Evans, T. Daniels, "Increasing Pilots Understanding of Future Automation State – An Evaluation of an Automation State and Trajectory Prediction System" Proceedings of the 38th AIAA/IEEE Digital Avionics Systems Conference, San Diego, CA, Sept 2019.

[17] S. G. Hart, L. E. Staveland, "Development of a NASA TLX (Task Load Index): Results of empirical and theoretical research," P. S. Hancock and N. Meshkati (Eds.), Human Mental Workload, Amsterdam: North Holland Press, pp. 139-183, 1988.

[18] S. G. Hart, "NASA-Task Load Index 20 Years Later," Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, pp. 904 – 908, Santa Monica: HFES, 2006.

[19] W. F. Moroney, D. W. Biers, F. T. Eggemeier, J. A. Mitchell, "A comparison of two scoring procedures with the NASA Task Load Index in a simulated flight task," Proceedings of the IEEE National Aerospace and Electronics Conference 2, Dayton, OH, pp. 734-740, 1992.

[20] S. J. Selcon and R. M. Taylor; "Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design," AGARD-CP-478, Situational Awareness in Aerospace Operations (pp. 5-1 to 5-8). Neuilly Sur Seine, France: Advisory Group Aerospace Research & Development (AD-A223939), April 1990.

[21] J. Brooke, "SUS – A quick and dirty usability scale," in P. W. Jordan, B. Thomas, B. A. Weerdmeester, A. L. McClelland, Usability Evaluation in Industry, London: Taylor and Francis, 1996.

[22] J. Brooke, "SUS: A Retrospective," Journal of Usability Studies, Vol. 8, Issue 2, pp. 29 – 40, Feb 2013.

[23] J. Sauro, A Practical Guide to the System Usability Scale, Denver: Measuring Usability, 2011.

[24] A. Bangor, P. Kortum, J. Miller, "Determining What SUS Scores Mean: Adding an Adjective Rating Scale," Journal of Usability Studies, Vol. 4, Issue 3, pp. 114 – 123, May 2009.