Evaluation of Pilot and Air Traffic Controller Use of Third Party Call Sign in Voice Communications with Pilot Utilization of Cockpit Display of Traffic Information

August 2015

Technical Report

Version: V_4
NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers’ names appear herein solely because they are considered essential to the objective of this report.
# Table of Contents

Executive Summary ..................................................................................................................... 7

1 Introduction ............................................................................................................................ 9

2 Background .......................................................................................................................... 10

3 Purpose .................................................................................................................................. 12

4 Objectives and Scope .......................................................................................................... 13

5 Test Support Requirements ............................................................................................... 14
  5.1 Resources/Logistics/Setup/Requirements ................................................. 14
  5.1.1 Subject Controllers ............................................................. 14
  5.1.2 ATC Lab Simulation ........................................................................ 14
  5.1.2.1 ATC Lab Operational Requirements ................................................. 14
  5.1.3 Subject Pilots ............................................................................................. 15
  5.1.4 Aircraft Simulator Operational Requirements .......................................... 15
  5.1.5 Video and Audio Recording ...................................................................... 16
  5.1.6 Human Factors, Pilot, and Controller Observers ...................................... 16

6 Test Procedure – Controller Specific Evaluation/Data Collection .............................. 17
  6.1 Assumptions .............................................................................................. 17
  6.2 Test Design ................................................................................................ 17
  6.2.1 Scenario Development .............................................................................. 17
  6.2.2 STARs and Sectors .................................................................................... 18
  6.3 Qualitative (Subjective/Observed) Data Collection ........................................ 18
  6.4 Quantitative Data Collection ..................................................................... 19
  6.5 Post-Evaluation De-brief ........................................................................... 20
  6.6 Observer Duties ......................................................................................... 20
  6.7 Human Factors Analysis (Air Traffic Controller Evaluation) ....................... 20
  6.7.1 Subjective Questionnaire Responses ......................................................... 21
  6.7.2 Direct Observation .................................................................................... 27
  6.7.3 Post Evaluation De-briefing Synopsis ....................................................... 27
  6.7.4 Subjective/Objective Conclusions (Controller Specific) ................................ 28

7 Test Procedure – Pilot Specific Evaluation/Data Collection ........................................ 29
  7.1 Test Assumptions ...................................................................................... 29
  7.2 Test Design .............................................................................................. 29
  7.2.1 Scenario Development .............................................................................. 30
  7.2.2 Pilot Scenarios ......................................................................................... 31
  7.2.3 Environmental Conditions ........................................................................ 32
  7.2.4 Simulator Configuration ............................................................................ 32
  7.3 Quantitative and Qualitative Data Collection .................................................. 32
7.3.1 Quantitative Data Collection ................................................................. 32
7.3.2 Qualitative Subjective and Observed Data Collection ......................... 33
7.3.3 Physiological Data Collection ............................................................... 34
7.4 Post Evaluation De-brief ......................................................................... 35
7.5 Observer/Evaluator Duties ...................................................................... 36
7.6 Human Factors Analysis (Pilot Evaluation) ............................................. 36
7.6.1 Quantitative Data Analysis ................................................................. 37
7.6.1.1 Pilot Performance (Errors) .............................................................. 37
7.6.1.2 Physiological Data Analysis .......................................................... 38
7.6.1.2.1 Eye Tracking .......................................................................... 38
7.6.2 Qualitative Data Analysis .................................................................... 44
7.6.2.1 Subjective Questionnaire Responses .............................................. 44
7.6.2.2 Direct Observation ........................................................................ 58
7.6.3 Post Evaluation De-briefing Synopsis ................................................ 58
7.6.4 Major Theme Discussion (e.g. Strategies, Crew Interaction, Fatigue, Situational Awareness, etc.) .......................................................... 59
7.6.5 Subjective/Objective Conclusions ....................................................... 59
8 Performance Data Analysis ........................................................................ 60
9 Conclusions ............................................................................................... 61
References ..................................................................................................... 62
Appendix A: Controller Test Anomaly Conditions and Schedule ...................... 63
Appendix B: Pilot Scenarios .......................................................................... 66
Appendix C: Controller Demographic Questionnaire ....................................... 72
Appendix D: Pilot Demographic Questionnaire ............................................ 73
Appendix E: Controller Post-Ren Questionnaire ........................................... 74
Appendix F: Pilot Post-Run Questionnaire ...................................................... 75
Appendix G: Post-Evaluation De-briefing – Controllers .................................. 76
Appendix H: Post Evaluation De-briefing – Pilots .......................................... 78

List of Figures
Figure 6-1: Controller Demographics .............................................................. 20
Figure 6-2: Mean Controller Responses TPCS vs TFID ................................ 22
Figure 6-3: Mean Controller Responses TPCS vs TFID (Hard/Medium/Easy) .... 23
Figure 6-4: Mean Controller Responses TFID Only .......................................... 24
Figure 6-5: Mean Controller Responses TPCS Only ......................................... 25
Figure 6-6: Mean Controller Responses TFID vs TPCS w/Max and Min .......... 26
Figure 7-1: Normal Line of Sight ................................................................. 35
Figure 7-2: Mean All Questions ....................................................................................... 48
Figure 7-3: A330 vs B737 Mean ...................................................................................... 49
Figure 7-4: Pilot Flying vs Pilot Monitoring Mean .......................................................... 50
Figure 7-5: Pilot Flying vs Pilot Monitoring, A330 vs B737 Mean ................................. 51
Figure 7-6: TPCS vs TFID Mean...................................................................................... 52
Figure 7-7: TPCS vs TFID Mean with Hi/Low ................................................................. 53
Figure 7-8: TPCS vs TFID, A330 vs B737............................................................... 54
Figure 7-9: TPCS vs TFID, Pilot Flying vs Pilot Monitoring Mean ............................... 55
Figure 7-10: Mean by Scenario Complexity .................................................................. 56
Figure 7-11: Pilot Flying vs Pilot Monitoring – Mean Complexity .............................. 57
Figure 7-12: TPCS vs TFID Mean Complexity ............................................................. 58

List of Tables
Table 7-1: Aircrew Demographics ............................................................................... 36
Table 7-2: Eye-Track Mapping Schedule ...................................................................... 39

List of Images
Image 7-1: Eye Track Mapping Reference Image (Boeing 737) ........................................ 35
Image 7-2: Heat Map A330 Captain Pilot Monitoring Easy ............................................ 40
Image 7-3: Areas of Interest A330 Captain Pilot Monitoring Easy .................................... 40
Image 7-4: Heat Map A330 Captain Pilot Flying Hard .................................................... 41
Image 7-5: Areas of Interest A330 Captain Pilot Flying Hard .......................................... 41
Image 7-6: Heat Map B737 Captain Pilot Monitoring Easy ........................................... 42
Image 7-7: Areas of Interest B737 Captain Pilot Monitoring Easy .................................. 42
Image 7-8: Heat Map B737 First Officer Pilot Flying Hard ............................................ 43
Image 7-9: Areas of Interest B737 First Officer Pilot Flying Hard .................................... 43
Executive Summary

In August 2007, the Surveillance and Broadcast Services (SBS) Program Office identified a potential programmatic risk to realizing all of the planned benefits that involved the use of voice-initiated Third Party Flight Identification (TFID) or Third Party Callsigns (TPCS). The risk was described as follows: “If the challenges with Air Traffic Control (ATC) use of third party aircraft flight ID in voice communications are not effectively resolved, the potential benefits of Interval Management and other future applications requiring the use of Cockpit Display of Traffic Information (CDTI) for spacing and separation of aircraft may not be realized.”

The purpose of this test was two-fold. Given data from previous phases of this evaluation indicating there may be a potential for pilot confusion by the aircraft being referred to during air traffic communication, this test was designed to validate whether that confusion happens and to what extent. Secondly, determine the impact of either of two phraseology formats; 1. Stating each individual letter and number digit (“Letters”) – TFID (e.g. Uniform Alpha Lima 123); 2. Stating the 3-Letter telephony call signs with reference material available to the pilot (“Telephonic”) – TPCS (e.g. United 123), each under the same flight conditions.

The objectives of this study were:
- Under varying test conditions, while manipulating flight profile complexity, to evaluate the potential for confusion when third party aircraft references are made in the National Airspace System (NAS)
- Evaluate the impact of either of two phraseology formats; TFID or TPCS
- Discern and evaluate controller and pilot perspective, respectively, from both a preference and usability (performance) standpoint.

The core assumption for the controller evaluation was: Call signs used would include only those that are currently used in the NAS (some were intuitive and typical; others were less obvious and rarely used); and for the pilot evaluation:
- Pilot experience/expertise with CDTI technology and simulation aircraft type was held in strict accordance with the test plan criteria;
- Call signs used included only those that are currently used in the NAS (some were intentionally intuitive and typical; others were less obvious and rarely used).

The pilot portion of this evaluation consisted of qualified and current line pilots from various airlines. Terminal Air Traffic duties were performed by trained and experienced air traffic controllers, provided by Flight Operations Simulation Branch (AFS-440), adhering to strict testing guidelines, designed to give each subject flight crew a consistent and like set of conditions. The ATC portion of the evaluation consisted of controller teams comprised of active and current terminal controllers. Pilots during this phase were qualified crews from AFS-440 and local offices.

An area of particular interest was pilot communication vigilance, as it specifically pertained to third-party radio communication. We had no direct way of determining if pilots were attending to and processing TPCS / TFID references, specifically. Pilots did give timely and accurate responses to all direct first-party transmissions, pointing to a
high degree of vigilance to radio traffic. A logical inference can be made that pilot vigilance to all third-party calls was equally high.

While it was beyond the scope of this evaluation to make a recommendation towards either format, the implementation of either one should take into account negative/positive habit transfer against current ATC culture and format. All that should be weighted against controller performance in this evaluation and expected performance in the future, when controllers will have the opportunity to train and reach a level of expertise that they currently enjoy with the formats that they currently use.

Despite controller preference for the callsign format over the flight ID format, subjective responses contradict that. All subjective response indices point to controller perception that the flight ID format is more effective from the standpoint of difficulty, timeliness and workload. This is the essence of the whole idea of controller preference versus actual usability of either or both formats.
1 Introduction

Aircraft equipped with Automatic Dependent Surveillance-Broadcast (ADS-B) Out will have the capability to transmit information (e.g., aircraft identification, track, altitude and speed) for presentation on graphical displays of other aircraft equipped with ADS-B In, CDTI, and ADS-B Guidance Display. The availability of this information in the cockpit fosters the development of multiple aircraft-to-aircraft applications supporting the Next Generation Air Transportation System Vision. It is initially envisioned that pilots and controllers will be able to use that information to enhance visual-based operations and situational awareness. These capabilities could afford the controller the opportunity to include the aircraft identification or call sign of a target aircraft in a traffic advisory transmission to another pilot for the purpose of unambiguous identification of the target or traffic aircraft on their display.
2 Background

In August 2007, the SBS Program Office identified a potential programmatic risk to realizing all of the planned benefits that involved the use of voice-initiated TFID or TPCS. The risk was described as follows: “If the challenges with controller’s use of third party aircraft flight ID in voice communications are not effectively resolved, the potential benefits of interval management and other future applications requiring the use of CDTI for spacing and separation of aircraft may not be realized.”

On June 30, 2010, the Federal Aviation Administration (FAA) chartered the ADS-B Aviation Rulemaking Committee, to provide a forum for the U.S. aviation community to define a strategy for incorporating ADS-B technologies into the NAS. One of the Aviation Rulemaking Committee’s recommendations was for the FAA to identify phraseology requirements, challenges, and risks associated with TFID. The recommendation further stated the FAA should form an appropriately supported action team to develop actual phraseology that can be validated through various human in the loop (HITL) analyses. To avoid confusion, validated phraseology should be used when a third party aircraft is referenced by its flight ID while in communications between own ship and ATC. This phraseology should allow pilots to distinguish between their aircraft when being issued instructions or communicated versus when they are being referenced by either ATC or another pilot.

The SBS Program Office formed an operational focus group (OFG), based, in part, on the recommendation of the ADS-B In Aviation Rulemaking Committee. The primary purpose of the OFG was to determine what, if any, new procedures or changes could be required to existing procedures. The OFG recommended four possible solutions, one of which was the existing phraseology as it was considered sufficient.

The OFG recommendations were incorporated into a TFID/TPCS HITL evaluation conducted by MITRE. The report (MTR130347R1) dated July 2013, evaluated the TFID/TPCS voice communication using the existing phraseology and three alternative candidates. It employed 88 pilots, 11 en route controllers, and 11 terminal controllers as participants.

The TFID solution validation plan is being conducted using a sequential, multi-phased approach with each subsequent phase building on the knowledge gained about the phraseology, terminology, procedures used, and mitigations developed for human performance issues discovered in the previous phases.

The FAA developed a TFID human factors analysis master test plan which included operational workshops and HITL simulations, broken out into three distinct phases. The first of the three phases (“phase 1”) was designed to gain input from both FAA and industry representatives to formulate candidate terminology that pilots and controllers consider viable for applications requiring TPCS. A group of operational experts concentrated on phraseology and terminology that was acceptable to controllers and provided unambiguous positive identification of target aircraft on a CDTI by the first party pilot. Based on the phase 1 report recommendations, three candidate solutions were tested in phase 2 simulation, of which two were deemed viable for final testing in phase 3.
Those two candidate solutions are explained below:

**Solution 1 - State each individual letter and number digit (“letters”) - TFID**

While using the existing ATC traffic phraseology as a baseline, add the identifier (as it appears on the ATC display and the CDTI traffic display) of the traffic being talked “about”. This relieves ambiguity between the controller and the pilot who must discern or correlate the three-letter identifier with the appropriate call sign (United). For example, “U A L 123” (uniform, alpha, lima, etc).

**Solution 2 - State the 3-letter telephony call signs with reference material available to the pilot (“telephonic”) - TPCS**

While using the existing ATC traffic phraseology, add the call sign of the traffic being referred to or talked “about”. The pilot will have access to 3-letter telephony call signs via electronic means or written reference. For example, “United 123”.
3 Purpose

The purpose of this test was two-fold. The data from previous phases of this evaluation, indicate a potential for pilot confusion when referred to, in a third party format, during air traffic communication. This test was designed to validate whether pilot confusion happens and to what extent. Secondarily, determine the impact of either of the two phraseology formats, each under the same flight conditions.
4 Objectives and Scope

- To evaluate the potential for confusion when third party aircraft references are made in the NAS, while varying test conditions, and flight profile complexity;
- To evaluate the impact of either of the two phraseology formats; TFID or TPCS; and
- To discern and evaluate controller and pilot perspective, respectively, from both a preference and usability (performance) standpoint.

NOTE: Pilot and controller studies will be conducted independently, but comments, discussion, and conclusions will be integrated.
5 Test Support Requirements
The setup of variables and scenarios to be tested were coordinated, developed, and programmed during the required pretest activities. The simulated airspace approximated that of Denver, Colorado and simulated in the AFS-440 ATC lab simulator.

The pilot portion of this evaluation consisted of qualified and current line pilots from various airlines. Terminal air traffic duties were performed by trained and experienced controllers, provided by AFS-440, adhering to strict testing guidelines, designed to give each subject flight crew a consistent and like set of conditions. The controller portion of the evaluation consisted of controller teams comprised of active and current terminal controllers. Pilots during this phase were qualified crews from AFS-440 and local offices. They performed duties in accordance with an established plan, designed to elicit real-world and representative responses from subject controllers. In most instances, pilot performance was directly affected by that of the controllers. For the test scenarios and conditions, see appendices A and B.

5.1 Resources/Logistics/Setup/Requirements

5.1.1 Subject Controllers
Subject controllers used for this data collection were required to be certified professional controllers, assigned to a facility level 10 or higher (i.e. Terminal Radar Approach Control Facilities equipped with the Standard Terminal Automation Replacement System (STARS) and Terminal Automation Modernization and Replacement Automation System).

The subject controllers were given a briefing at the beginning of the session. The briefing included information that pertained to this specific operation and airspace. Prior to testing, subject controllers were asked to complete a demographic questionnaire, see appendix C. A post-session questionnaire was given to each subject after completion of each scenario, see appendix E. Upon completion of all scenarios, a post-simulation questionnaire and de-briefing was conducted by test personnel, see appendix G.

5.1.2 ATC Lab Simulation

5.1.2.1 ATC Lab Operational Requirements
The AFS-440 simulators are comprised of three separate simulator systems; AIRBUS 330/340 Level D qualified simulator, BOEING 737-800 Level D qualified simulator, and an ATC lab simulator. All three simulators can be run independently or connected together with high level architecture software protocols.

The simulator was configured to utilize runways 34R, 35L, and 35R at a generic airport.

The airspace was configured with two operational sectors (feeder and final), seven STARS, one over-flight route, and three Instrument Landing System (ILS)s. The subject feeder controller was required to vector aircraft that were arriving from the STARS to align for in-trail spacing. The feeder controller handed off aircraft to the final controller for vectoring to an ILS approach. The controllers were required to use separation requirements in accordance with Order 7110.65 and ADS-B separation. [1]
The test scenario controller work stations utilized a STARS color digital display with fusion capability.

5.1.3 Subject Pilots

Subject pilots were required to be current line pilots from various airlines, qualified in the Boeing 737, and Airbus 330 aircraft type, respectively. It was highly desired that “talked to” subject pilots have more than a cursory background in the use of ADS-B and CDTI. It was critical that pilots have the necessary test qualifications and demographic background that in the case of a pilot absence, the test director was prepared to make the critical decision to postpone or terminate that day’s particular data-collection.

A pre-brief was conducted with the aircrew and the observers covering the test requirements and subject responsibilities. The briefing included operational requirements as they pertained to this specific evaluation. Approach charts were provided to the pilots, either in paper format or electronically. Also, prior to testing, subject pilots were asked to complete a demographic questionnaire, see appendix D.

A post-scenario questionnaire was given to each subject after completion of each scenario, see appendix F. Upon completion of all scenarios, a post-simulation questionnaire and de-briefing was conducted by test personnel, see appendix H.

5.1.4 Aircraft Simulator Operational Requirements

The following operational requirements applied to all data collections involving the AIRBUS 330/340 Level D simulator, BOEING 737-800 Level D:

- Auto-pilot and auto-throttles will be operational and operated in accordance with aircrew company policy;
- Simulator aircraft will provide suitable guidance for hand flown operations using the flight director, as well as fully coupled autopilot operations involving ILS instrument approaches and missed approach procedures;
- Simulator aircraft will conform to Level D qualifications;
- Simulator aircraft will be configured to conduct high level architecture at the highest design fidelity;
- Simulator aircraft will be configured to conduct the FRNCH TWO arrival at Denver International Airport;
- Simulator aircraft will have fully functional voice communications capability to include, at a minimum three headsets. One headset for each pilot and one for an observer;
- The A330 simulator will have ADS-B, CDTI, and ADS-B guidance display fully functional;
- The simulator will have no artificial errors introduced;
- Simulator visual systems must be operating at the highest design fidelity; and
- All simulated traffic should appear visually (out-the-window) at a minimum of 15 statute miles.
5.1.5 Video and Audio Recording

Video and audio recordings were made of each simulator session, commencing upon subject pilot entry into the cockpit and ending no earlier than subject pilots departing the cockpit.

Video recordings were made of both the left and right forward panel targeting specifically each navigation display and also the left and right electronic flight bags, respectively.

The requirement to record the audio and video of the simulator was briefed to the participants before commencement of the data collection.

NOTE: No video was taken within the ATC lab simulator during the controller testing event(s).

5.1.6 Human Factors, Pilot, and Controller Observers

An instructor operating station (IOS) operator, a human factor observer, and a pilot observer were unobtrusively positioned inside the simulator, behind the pilot stations for each respective pilot evaluation. The human factor observer was a qualified human factor specialist, capable of observing, interpreting, and capturing the essential elements of individual pilot and crew interaction as it directly affects performance.

Similarly, a human factor observer and two controller observers were positioned behind the subject controllers’ work station. Each was capable of observing, interpreting, and capturing essential elements of individual and collective controller interaction as it affects performance.
6 Test Procedure – Controller Specific Evaluation/Data Collection

6.1 Assumptions
Call signs used would include only those that are currently used in the NAS (some were intuitive and typical; others were less obvious and rarely used).

6.2 Test Design
Controllers were evaluated in pairs. Each subject controller pair was evaluated in two-hour sessions. During each session, the controllers were required to use one of the two phraseology options for the entirety of that session. Controllers used the other of the two options in the second session.

Prior to the first session of each controller pair, controllers were given an in-brief. Following each session, each pair was given a post-session questionnaire, followed by a break. During that time, the next subject pair was given their in-brief, followed by a two-hour evaluation. This rotation continued through four total sessions (two for each respective pair).

NOTE: Depending on time, controllers were given a short tour of the cockpit and relevant displays (i.e. CDTI and ADS-B guidance display). This occurred either before or after the controller evaluation. If time permitted, pictorial representations of relevant cockpit display information were shown during the in-brief.

During the course of their duties in this data collection, controllers may or may not have actually decided to use “Third Party” transmissions. For that reason, evaluators directed controllers to artificially make a “Third Party” transmission any time a CDTI-equipped aircraft passed by DORKE, BABAA, and HIMOM intersections. CDTI-equipped aircraft were designated by an “open” dot in its respective data-block. This helped to ensure that “Third Party” transmissions were made and appropriately evaluated.

6.2.1 Scenario Development
During each session, the following is a list of independent variables (conditions) that were to be manipulated in order to elicit varying levels of activity (workload):

- Traffic density;
- Incorrect pilot read-back;
- Weather (necessitating a traffic deviation);
- Pilot talked about accepting clearance for the talked to aircraft (use same company ID and a similar flight number, e.g., UAL 151 and UAL 191) and taking immediate action on that clearance such as a turn or climb;
- Pilot unable to accept clearance for whatever reason;
- Pilot not acknowledging a clearance;
- Go-around; and
- Balk landing.
Given the relatively limited number of sessions and short duration of each session, test conditions were set up using an “easy/medium/hard” methodology. Controllers were evaluated based upon their performance as the conditions changed in each session. The controller evaluation followed the methodology below:

- Each scenario was comprised of aircraft arriving in a continuous flow on a designated STAR;
- Aircraft were vectored for in-trail spacing to a designated ILS approach;
- In-trail spacing (miles-in-trail) conformed to existing standards;
- The flights were a mix of commercial and general aviation aircraft;
- In all scenarios, callsigns were a mix of valid, non-intuitive (infrequently used), common domestic, and international;
- For this data collection effort, the simulated aircraft flew the published STAR and overflight routes, adhering to all speed and altitude constraints;
- Departures occurred, flying preferential departure routes;
- A Fusion display with ADS-B capability was used; and
- Controllers were permitted to use ADS-B for separation.

6.2.2 STARs and Sectors

The following STARs were utilized:

- AYIYU (fictitious);
- MOLTN;
- FRNCH;
- HEPEG (fictitious);
- LDORA;
- BOSSS;
- KOHOE; and
- Overflight (unnamed).

Sectors:

- Feeder; and
- Final.

After completion of each session, the electronic data and de-briefing data was assembled for post-test statistical analysis. Approximately four hours total time (divided equally between both TFID and TPCS formats) was required for collecting data from each pair of controllers and conducting the de-briefing.

Each session contained three scenarios and ran for approximately four hours. The current equipment setup allowed two controllers to participate in each session. The sessions were scheduled each day utilizing similar easy/medium/hard scenario formats.

6.3 Qualitative (Subjective/Observed) Data Collection

Human factor and engineering psychology analysis was included as part of this evaluation. After each run, the human factor observer guided the subject controllers through a post-session questionnaire, to be completed while the evaluation team re-configures for the next session. After completion of the entire evaluation, the lead human factors observer guided the subject controllers through the post-simulation de-brief.
For post-run and post evaluation questionnaires, see appendices E and G, respectively. A multi-dimensional rating procedure is used. These questions solicited controller feedback on their perception of performance, workload, and comfort. Given the intrusive nature of any data-gathering procedure of this type, we minimized the number of questions and the time required to complete the questions.

Objective controller performance measures were limited in scope. This was primarily accomplished through simple observation of controller performance. Observation data was taken by an ATC subject matter expert, or human factor subject matter expert, each of whom was positioned behind the subject controllers. Each of the observers were retired professional controllers with multiple years of controller experience and previous experience as observers. All scenarios were carefully scripted. During those periods in a given ATC sequence, when a controller(s) might have performed out of the norm from what was either expected or planned, both primary and secondary task completion were monitored. The basic logic behind the primary and secondary task measure methodology is that spare mental and physical capacity, not being used by the primary task, will be devoted to accomplishing the secondary task. Observers were primarily concerned with controller comfort, effectiveness, and responsiveness with each phraseology option.

Observers culled through and organized their notes, identifying major themes, issues, problems, etc. from which a discussion could be made and conclusions drawn. From there, a connection to subjective, physiological, and aircraft performance may arise and a chain of causation may be established, pointing directly to the impact of the third party communications on performance. That discussion is included later in this document.

6.4 Quantitative Data Collection

As in the pilot portion of this data collection, quantitative data was elicited as a result of first party and third party transmissions. That data was presented to the Flight Systems Laboratory Branch (AFS-450) in .csv, .txt, or .xls file format.

Within the letter (TFID) and telephonic (TPCS) formats, respectively, the number of third party transmissions were varied. That number was commensurate with the level of difficulty within each category and fell within a range, the actual number of which was established during the test.

NOTE: The total number of transmissions within each format was intended to be as close to the same for each format.

Range of potential events:

- Easy = 1 – 2
- Medium = 3 – 4
- Hard = 5

The targeted quantitative data was a function of the time of each transmission per each individual controller and mean total transmission time of all controllers. The intent was to evaluate latency of controller reaction (the total time from onset of an event to the actual depression of the press-to-talk button immediately before the controller transmits) with each of the required phraseology options.
6.5 Post-Evaluation De-brief
After completion of the simulator session, all data collection participants conducted a de-briefing. The lead human factor observer led the de-briefing. Discussion covered data collection execution, review of post scenario questionnaires and any problems that arose during the data collection.

6.6 Observer Duties
All observers, including the IOS operator, annotated events, activities, conditions, actions, and communications that the respective observer felt to be significant to the data collected. Observers submitted observer notes in electronic format.

6.7 Human Factors Analysis (Air Traffic Controller Evaluation)
On April 21st, 2014, all evaluation team members conducted briefings and performed critical path checks, insuring that all software, hardware, and personnel were ready. The controller portion of the TFID evaluation was conducted on April 22nd to 24th 2014. On each day, two pairs of certified National Air Traffic Controllers Association controllers were evaluated; two in the morning session and two in the afternoon session. Controller demographic information is presented below.

![Controller Demographics Diagram](image)

Prior to the start of each evaluation session, the test director and lead controller evaluator presented a briefing to the participants. The briefing consisted of background information, figure representations of the test controller environment, and phraseology options the subjects would be required to use. Each controller pair was given both easy, medium, and hard scenarios in each of the third party callsign and third party flight ID phraseology formats. In addition, each controller alternately served in both the feeder and final controller positions. With the exception of one controller team, all teams...
participated in 6 x 20-30 minute scenarios. After each session, the controllers were required to answer a 5-question questionnaire, designed to gather controller perception of performance, comfort, and workload.

Prior to the actual data collection, the team of evaluators established a list of anomalies that would be presented, individually, to one or both controllers at appropriate and realistic times during the simulation. Ideally, the schedule of these anomalies would be identical across all subjects. The fluid nature of the simulation did not allow this since all subject controllers did not experience identical traffic flows and conditions. For the list of anomalies and presentation schedule, see appendix A. Note that the total number and types of anomalies, per subject controller, was fairly constant. The purpose of the anomalies was to evaluate controller performance when presented with situations considered out-of-the-norm. In essence, evaluators induced activity levels that may or may not have affected controller performance or workload. Controller observers recorded anomaly onsets and resultant controller reaction (e.g. anomaly recognition, potential response errors, latency of response, etc.). The anomalies (induced by confederate pilots) are listed below, along with the frequency of their occurrence:

- Incorrect readback (40);
- Not in site (24); blocked transmission (XMNT)(17);
- Was that for me?(19);
- Debate instructions(21);
- Flip the switch(10);
- ATIS message missing(12); and
- Other (6).

6.7.1 Subjective Questionnaire Responses

When given instructions for the completion of each run’s questionnaire, subject controllers were asked to base their responses upon relative comparison to their “normal” air traffic activities.

The tables and figures associated with this portion of the report were organized in accordance with the following convention: horizontal versus vertical bar alignment; ordinate scale; focused variables: mean values all; mean values by TFID vs TPCS and scenario complexity, see figures 2 to 5.

Overall, irrespective of induced run difficulty, subject controllers indicated that the use of third party callsign phraseology was easier to use and resulted in greater perceived comfort and less workload. The subjective responses also indicated that controllers felt that the TPCS phraseology enabled them to be timelier (faster). When queried specifically about this during the post-session de-briefing, controller responses indicated that the TPCS format is what they were accustomed to using in the scope of their current duties. When they encountered a callsign that they were not accustomed to using habitually, the easier (faster) format was the TFID.
Figure 6-2: Mean Controller Responses TPCS vs TFID
Figure 6-3: Mean Controller Responses TPCS vs TFID (Hard/Medium/Easy)
Figure 6-4: Mean Controller Responses TFID Only

Mean Subjective Controller Responses - Third Party Flight ID Only (Compared to Normal Operations)

- Difficulty
  - Much More Difficult
  - Very Uncomfortable
  - Much Slower
  - Much Higher

- Comfort
  - Much Easier
  - Very Comfortable
  - Much Faster

- Timeliness
  - Same as Typical Operation
    - Much Lower

- Individual Workload
  - Much Lower

- Collective Workload
  - Much Lower

TFID
Issued on [August 2015]
[Flight Operations Simulation Branch, AFS-440]
Figure 6-5: Mean Controller Responses TPCS Only

Mean Subjective Controller Responses - Third Party Callsign Only (Compared to Normal Operations)

- TPCS Easy
- TPCS Medium
- TPCS Hard

- Much More Difficult
- Very Uncomfortable
- Much Slower
- Much Higher
- Much Higher
- Much Higher
- Much Slower
- Much Higher
- Much Higher

- Very Comfortable
- Much Faster
- Same as Typical Operation
- More Difficult
- Much Slower
- Much Higher
- Much Higher

- Difficulty
- Comfort
- Timeliness
- Individual Workload
- Collective Workload

TFID
Issued on [August 2015]
[Flight Operations Simulation Branch, AFS-440]
Figure 6-6: Mean Controller Responses TFID vs TPCS w/Max and Min
6.7.2 Direct Observation

Controller-specific activities – the evaluation was intentionally setup, using a mix of common, frequently used callsigns, and less common, obscure, and counter-intuitive callsigns. Although, a sheet with all callsigns was available at the subject controllers’ workstation, they seldom referred to it and when not sure of the callsign presented on the radar, instinctively reverted to the alpha numeric (TFID) callsign. This presents anecdotal evidence that the TFID format may be more intuitive than the TPCS format, despite subjective controller feedback that would suggest otherwise.

Concerning human factors specific activities: there appears to be a negative habit transfer issue with the phraseology that controllers normally use that’s in accordance with their training, experience, and requirements. When workload/traffic and density/activity levels are higher (i.e. when controllers rely on a certain level of “fast” or automatic response), a natural tendency is to use what comes naturally, with less cognition required.

A potential source of interference with controller habits and expectations is the use of the “dot” in the data-block which, depending upon whether it is solid or open, indicates whether that specific aircraft is equipped with a CDTI and ADS-B. That information impacts controller strategy and phraseology formulation. While this was certainly a training/learning artifact, no conclusions may be drawn about the cause and effect of this symbology. Several times throughout each of the sessions, controllers were subtly reminded about the dot’s intent as well as the phraseology format they were to use.

6.7.3 Post Evaluation De-briefing Synopsis

During the course of each de-briefing, several consistent comments and themes were voiced:

Of the 11 subject controllers who participated in this study, 9 preferred the callsign (TPCS) phraseology format, since that most closely associated with what they currently use. Logically, there would be less negative habit transfer between this format and what they were culturally accustomed to using. As a caveat, several controllers did point out that the callsign format may have caused some confusion, especially when infrequently used or non-intuitive callsigns were used. The two controllers that preferred the flight ID format, indicated that it eliminated any confusion about the reference aircraft since that information could be effectively validated by pilots via CDTI data in the cockpit.

This leads to a discussion of the next comment, one that pertains to the amount of information controllers would pass along if required to voice either or both formats.

More than half the participants felt a need to give more information to the pilots when referring them to another aircraft (e.g. aircraft type, altitude, direction, etc.). With either of the test phraseology formats, controllers were directed only to give aircraft callsign when referring to a third party. This is a departure from current ATC phraseology norms and may lead to negative habit transfer. Additionally, a like number of subjects would prefer having the option to add additional information as they deemed necessary, despite having been briefed that pilots would have access to all that data via the CDTI.
As a suggestion, a few controllers offered that the callsign format would be optimal, but could be augmented with using the flight ID format at their discretion, should they need to. This might happen during those times when an infrequent, non-intuitive callsign was portrayed (e.g. cactus (AWE), waterski (LOF), ETC).

One confusing aspect of the radar data-block was the use of the “circle” (∙) symbol, preceding each aircraft callsign. Subject controllers voiced that it was initially confusing, having to determine whether the presence of the “circle” represented an aircraft that was or was not equipped with ADS-B technology. The presence of a filled circle means the aircraft is not ADS-B equipped. NOTE: Only ADS-B "OUT" equipped aircraft will transmit their flight ID and only ADS-B "IN" aircraft will have the capability to display that same information via a CDTI.

Given the unfamiliar nature of what the “circle” portrayed, cognitive resources were used, albeit fleetingly, to verify its meaning, establish a strategy, and execute a voice command.

6.7.4 Subjective/Objective Conclusions (Controller Specific)

In making any decision to engage either or both format options, much credence has to be given to controller preference, remaining cognizant of what they are currently accustomed to using. While it’s beyond the scope of this evaluation to make a recommendation towards either formats, the implementation of which should take into account negative and positive habit transfer against current ATC culture and format. All that should be weighted against controller performance in this evaluation and expected performance in the future, when controllers will have the opportunity to train and reach a level of expertise that they currently enjoy.

Despite controller preference for the callsign format over the flight ID format, subjective responses contradict that. All subjective response indices point to controller perception that the flight ID format is more effective from the standpoint of difficulty, timeliness, and workload. This is the essence of the whole idea of controller preference versus actual usability of either or both formats.
7 Test Procedure – Pilot Specific Evaluation/Data Collection

7.1 Test Assumptions

- Pilot experience and expertise with CDTI technology and simulation aircraft type will hold in strict accordance with section 5.1.3 above; and
- Call signs to be used will include only those that are currently used in the NAS (some will be intuitive and typical; others will be less obvious and rarely used).

7.2 Test Design

The following section describes the technical conditions, procedures, and requirements for the TFID, with particular focus on aspects applicable during the flight simulation and data collection phase. This evaluation focused almost exclusively on the HITL with limited corroborating performance data in the form of simulator data output. As previously stated, we were concerned with the performance of primary aircrews as they are “talked to”, third party aircraft as they are “talked about”, and air traffic controllers’ effectiveness with conveying essential information with the two phraseology options. To effectively evaluate pilot and controller performance, this data collection and analysis was broken down into three distinct phases, while independently isolating each of the above-mentioned subject populations. During each evaluation phase, as one subject population was being evaluated (i.e. performance monitored and analyzed as a dependent variable), the actions of the other two populations were held constant and scripted, with the intent of providing the same input and stimuli to the evaluated crew and controllers across all subjects in that group. This will allow for consistent post data-collection analysis later on (i.e. isolation of cause-and-effect). However, we were able to extract additional “talked to” and “talked about” data during some test phases when the focus was not “talked to” or “talked about” respectively.

Test scenarios were designed within each of three categories, easy (“low complexity”), medium (“medium complexity”), and hard (“high complexity”). Each category was designed to hold constant those variables that would directly influence pilot activity and performance. They will be discussed later in this document. The intent was to manipulate procedural requirements of subject pilots and controllers to induce varying levels of workload. During those periods, in a given flight sequence when a pilot and crew might have had to perform a function out of the norm from what was either expected or planned, both primary and secondary task completion were monitored. The basic logic behind the primary and secondary task measure methodology is that spare mental and physical capacity, not being used by the primary task, will be devoted to accomplishing the secondary task. In this case, the primary task was the performance of approaches under the specified conditions of this evaluation. The secondary tasks were aircraft control within stabilized flight criteria, communication procedures, timely checklist completion, etc. The greater the demand for resources made by primary tasks the less resources are available for performance of secondary tasks. Specifically, during periods of heightened activity or workload, reaction times, latency of task completion or task shedding may have taken place. These events were observed, recorded, and analyzed, commensurate with aircraft performance metrics.

This data collection effort was conducted over 6, 4-hour simulator periods in the Boeing 737 and Airbus 330 Level D, full flight simulators. Each aircraft will be evaluated in separate and distinct periods.
A pilot in-brief preceded the initial simulator session which was divided into two, two hour simulator blocks. At the beginning of the first simulator session of the day, each pilot flew a warm-up approach. There was one, 15 minute break at the mid-point of the data collection. Pilots were instructed to fly and configure their aircraft in accordance with current company standard operating procedures and techniques.

After each distinct session (i.e. low, medium, and high complexity), each subject completed a short computer-based questionnaire, designed to elicit their perceived sense of comfort, workload, and performance. For the pilot questionnaires, see appendix F. At the end of the entire evaluation, pilots were given a short break and then participated in the post-simulation de-brief, see appendix H.

7.2.1 Scenario Development

Test administrators manipulated the following variables: Weather conditions, traffic flow, communication density, and frequency of radio calls that direct changes to aircraft flight profiles. Variable mixes within each group are listed below.

NOTE: In all scenarios, callsigns were valid, non-intuitive, and infrequently used callsigns.
**Easy (Low Complexity):**
- Day, visual meteorological conditions, and smooth air;
- Very low radio “chatter”;
- Frequency changes;
- Three – six TFID traffic call outs;
- Low aircraft count;
- Follow aircraft on arrival;
- Be issued and fly a CDTI Assisted Visual Separation (CAVS) approach behind this aircraft; and
- Minimal “talked about” during the “talked to” specific evaluation.

**Medium (Medium Complexity):**
- Night, visual meteorological conditions, and light turbulence (40%);
- Moderate radio chatter, multiple runway ops;
- Frequency changes;
- Four – seven TFID traffic call outs;
- Medium aircraft count;
- Follow aircraft on arrival;
- Be issued and fly a CAVS approach behind this aircraft; and
- Moderate “talked about” during the “talked to” specific evaluation.

**Hard (High Complexity):**
- Night, instrument meteorological conditions, moderate turbulence (80%), convective activity around area, ceiling, and visibility 4,000-5;
- Heavy radio chatter, multiple runways;
- Five – seven TFID traffic call outs;
- High aircraft count;
- Thunderstorm on arrival between DORKE and ARCHY;
- Runway or approach change;
- Follow aircraft on arrival;
- Be issued and fly a CAVS approach behind this aircraft if able;
- Call from flight attendant in back; and
- Numerous “talked about” during the “talked to” specific evaluation.

### 7.2.2 Pilot Scenarios

A total of three scenarios were used under the following guidelines:
- Each scenario can be used in either a “talked to” or “talked about” environment (easy – low complexity, medium – medium complexity, and hard – high complexity); and
- Each crew will perform all three scenarios with the main focus on either the “talked to” category or “talked about” category using one of the two communication options (NOTE: as previously discussed, to enhance realism, additional “talked to” and “talked about” data may be collected although it may not be the main focus). Crews will take a break and repeat the scenarios using the other of the two communication options.
7.2.3 Environmental Conditions

The following environmental conditions were manipulated at specified points during collection sessions:

- Ceiling;
- Day and night;
- Visibility;
- Turbulence level; and
- Winds.

For specific scenarios, see appendix B.

7.2.4 Simulator Configuration

The simulator was configured prior to each run as follows:

- Over FRNCH on FRNCH TWO arrival;
- Altitude – FL250;
- Speed – 280 knots indicated airspeed (KIAS);
- Configuration – clean;
- Flight Management System (FMS) configured for the standard terminal arrival;
- Very high frequency radios set to the appropriate approach control and localizer navigation frequencies;
- All airfield approach lighting systems operate normally (set to level 5); and
- Aircraft weight and balance set to medium weight. Presets below:
  - B737 GW = 130,000#; ZFW=110,000#; Fuel Weight=20,000#; CG=26%
  - A330 GW = 352,740#; ZFW=297,630#; Fuel Weight= 55,115# ZFCG=33.7%

7.3 Quantitative and Qualitative Data Collection

7.3.1 Quantitative Data Collection

All quantitative data was elicited by first party and third party transmissions. That data was presented to Flight Systems Laboratory Branch (AFS-450) in .csv, .txt, or .xls file format.

For both the telephonic and letter format, the number of third party transmissions varied. That number was commensurate with the level of complexity within each category and fell within a range, the actual number of which was established during the test.

Range of potential events:

- Easy (low complexity) = 1 – 2
- Medium (medium complexity) = 3 – 4
- Hard (high complexity) = 5
Additionally, ATSI (data collection, reduction, and analysis contract support) provided position data of all ATC lab simulator aircraft that were actively "flying". This position data was collected at a rate of at least 5Hz. The data included at least the following items:

- A time reference synced with B737-800 and A330 simulators;
- Aircraft latitude;
- Aircraft longitude;
- Aircraft altitude;
- Communications from the subject pilots to capture any confusion along with corresponding observer event marks; and
- Any other aircraft state vector data items that might be available. The examples are below:
  - Aircraft ground speed;
  - Aircraft deviation from vertical track to be flown or altitude to be maintained;
  - Aircraft deviation from lateral course or vector to be flown; and
  - Aircraft ground track.

ATSI also provided data from discrete inputs that are a result of human intervention (controllers and/or simulator operator and pseudo pilots). The data included the following items:

- Time in which units are in sync with B737-800 and A330 simulators, for controllers pressing the push to talk button;
- Time in which units are in sync with B737-800 and A330 simulators, for controllers releasing the push to talk button; and
- Corresponding .wav files of controllers callouts during push to talk transmission.

### 7.3.2 Qualitative Subjective and Observed Data Collection

Three observers were present in the cockpit throughout the entirety of the pilot portion of the test. One was a current commercial pilot, well-versed in the techniques and procedures of both simulators; one was a retired commercial pilot, versed in general air carrier techniques and procedures for both simulators, but also an expert in the field of human factors engineering; one was a human factors engineer, serving as the primary evaluator.

Human factor and engineering psychology analysis is included as part of this evaluation. After each run, the human factor observer guided the subject pilots through a post-run questionnaire, to be completed while the simulator operator re-configured for the next run. After completion of a full data collection, the test director or his representative will guided the subject pilots through the post-simulation de-brief.

For post run and post data collection questionnaires, see appendices F and H. A multi-dimensional rating procedure was used. These questions solicited pilot feedback on their perception of stabilized flight, visual acquisition, workload, and comfort. Given the intrusive nature of any data-gathering procedure, of this type, we minimized the number of questions and the time required to complete the questions.

Objective crew performance measures were limited in scope. This was accomplished through simple observation of pilot and crew performance. Observation data was taken by in-the-cockpit observers (both pilot and human factor specific). All flight scenarios were carefully scripted.
During those periods in a given flight sequence when a pilot and crew might have had to perform a function out of the norm from what is either expected or planned, both primary and secondary task completion were monitored. Specifically, during periods of heightened activity or workload, reaction times, latency of task completion or task shedding may have taken place. When observed, those events were recorded, and analyzed commensurate with aircraft performance metrics.

Examples of such events are listed below:

- Missed radio calls;
- Query controller for a clearance repeat;
- Inappropriate response to a radio call to another aircraft;
- Misunderstood clearance and/or corresponding incorrect reaction;
- Incorrect aircraft tag on CDTI;
- Error in the use of aircraft Flight Management Computer and Navigation System;
- Missed or incomplete checklist;
- Latency in radio response; and
- Airspeed, altitude, or course deviation.

Observers culled through and organized their notes, identifying major themes, issues, problems, etc. from which the analysis and discussion was generated. Any possible connection to subjective, physiological, and aircraft performance was made, establishing a potential chain of causation, pointing directly to the impact of the third party communications on performance.

### 7.3.3 Physiological Data Collection

AFS-440 has the capability to provide a human interface in the form of non-invasive video-based glasses with audio recording capability. The device is worn like a normal pair of glasses and includes a high-definition scene camera and special eye-tracking technology that captures the eye movement of the participant(s) wearing it. The eye-trackers can be used to record subject pilot point-of-regard, saccade rates, dwell time, head movement, and blink rate, all potentially correlates to workload, task efficiency, and deficiency. After data-collection analysis, eye fixations, measured at 80ms or greater, are mapped onto a reference image (e.g. Airbus/Boeing cockpit) for appropriate analysis. An example of the Boeing 737 reference image, taken from the first officer side, is below, see image 7-1. The reference image was created from 10 separate images, merged into one clear, comprehensive representation. The reference image is not to scale and is deliberately distorted to accommodate all instruments and controls in one single picture. This facilitates mapping of all scans/saccades during the entirety of each run.

NOTE: Eye tracking technology was not used during the controller portion of the study.

A pilot’s primary field of view (FOV) is defined as the area where the pilot should be able to use all the required instruments with “minimum head and eye movement". This area is depicted inside the red circle on image 7-1. Primary optimum FOV is based on the vertical and horizontal visual fields from the design eye reference point (i.e. a single reference point in space selected by the designer where the midpoint between the pilot’s eyes is assumed to be located when the pilot is properly seated at the pilot’s station that can be accommodated with eye rotation only).
With the normal line-of-sight established at 15° below the horizontal plane, the values for the vertical and horizontal (relative to normal line-of-sight forward of the aircraft) are ±15°, as shown in figure 7-1. This area is normally reserved for primary flight information and high priority alerts. Naturally, information in this area will be detected more quickly than if placed outside this area.

Image 7-1: Eye Track Mapping Reference Image (Boeing 737)

Figure 7-1: Normal Line of Sight

7.4 Post Evaluation De-brief
After completion of the simulator session, all data collection participants participated in a de-brief of the subject pilots and controllers. The test director or human factor observer led the de-brief. Discussion covered data collection execution, review of post approach questionnaires
and any problems that arose during the data collection. The lead human factor observer also administered the post data collection questionnaire, see appendix H.

7.5 Observer/Evaluator Duties

All observers/evaluators, including the IOS operator, annotated events, activities, conditions, actions, and communications that he may have felt was significant to the data collection. As noted above, observer and evaluators individually determined the significance of any event(s) that warranted particular note. Observers transferred notes to an observer notes form, which was a modified data collection condition matrix (provided by test director) and submitted the observer notes form in electronic format.

7.6 Human Factors Analysis (Pilot Evaluation)

The pilot portion of the evaluation took place from July 21\textsuperscript{st} through August 1\textsuperscript{st}, 2014. From July 14\textsuperscript{th} through July 18\textsuperscript{th}, 2014, all evaluation team members conducted briefings and performed critical path checks, insuring that all software, hardware, and personnel were ready. On each day, a full crew (captain/first officer) was evaluated. Over the course of two weeks, the team evaluated 10 separate crews. Six crews were evaluated as subjects in the Airbus 330-200 (A330), and four crews were evaluated in the Boeing 737-800 (B737) Level D full flight simulators. The crews were qualified and current line pilots, with each crew from the same air carrier, see table 7-1. Four different air carriers were represented. NOTE: On three occasions, a first officer or captain subject pilot was not able to attend the evaluation sessions. In each case, an evaluation team member performed duties as the missing crewmember. Note, it was the same pilot for each of the three occasions. He was current and qualified in the type aircraft, although he was not completely familiar with the company techniques and procedures of the subject crewmember with whom he was flying. He was instructed to perform as he normally would on the line, not divulging any of the specifics of the test beyond what was briefed prior to performing. Any performance impact will be discussed further in this report. An ATC lab simulator was integrated with both full flight simulators. Lab personnel consisted of retired air traffic controllers, aircraft pseudo pilots, and support engineers.

<table>
<thead>
<tr>
<th>Table 7-1: Aircrew Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crew #</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

The evaluation team consisted of a test director, pilot observer, and two human factor observers. Several industry subject matter experts also observed from the ATC lab simulator. Prior to the start of each evaluation session, the test director and lead human factors observer presented a
briefing to the participants. The briefing consisted of background information, figure representations of the test environment, flight profiles to include STAR and APPROACH charts and a brief reference to NEXTGEN phraseology options the subjects would encounter. Each session lasted approximately three hours and was followed by a de-briefing, conducted by the lead human factors observer, with participation from evaluation members and the subject pilots.

Each subject crew was required to fly six distinct scenarios, each lasting approximately 25 minutes. Across those six scenarios, complexity of flight profiles was manipulated to elicit variable crew performance within the context of the stated purpose of the evaluation. We were aware of potential primary and secondary task shedding. Two scenarios were considered relatively benign and straight-forward in accordance with typical pilot/crew duties; two scenarios were moderately complex and two scenarios were configured to be still more complex. After each scenario, the subject pilots answered a five question questionnaire, designed to gather perception of performance, comfort, and workload.

Prior to the actual data collection, the team of evaluators established a list of data-collection points when “talked about” events would be presented throughout each scenario. These points were inserted at appropriate and realistic times during each profile. Ideally, the schedule of these events would be identical across all subjects. Given the fluid and variable nature of the simulation, not all crews experiences all events at exactly the same time nor under exactly the same conditions. They did occur within a close enough tolerance of each other, commensurate with the realism of the scenarios, such that crews could be evaluated across those same variables. As such, the total number and types of events, per subject crew, were fairly constant. Any significant variations from planned events were noted. The reason for those events were to evaluate individual and crew performance when presented with varying levels of activity, in concert with proposed third party communication references. In essence, evaluators induced activity levels that may or may not have affected pilots’ performance or workload. Pilot and human factors observers recorded subject pilot reaction to pre-established data collection events. (e.g. event recognition, potential response errors, latency of response, etc.).

All scenarios and data collection events were carefully “scripted” and vetted, well in advance of the evaluation. They were choreographed during each session by the test director and coordinator in the ATC lab simulator and coordinated with observers in the cockpit.

Note: A330 crews 2, 3, and 6 were from the only air carrier with operational CDTI technology. These crews characterized their training as minimal and experience low. Crews characterized their company’s use of CDTI technology, as being in an introductory phase, with no operational implementation. The remaining 2, A330 crews and the 4, B737 crews were from companies that did not have CDTI technology. For test purposes, their units were turned off.

7.6.1 Quantitative Data Analysis
7.6.1.1 Pilot Performance (Errors)
There were 60 data collection runs, over the 10 day data collection effort. The 60 runs resulted in the subject pilots hearing approximately 18,280 total radio transmissions, of which approximately 9,460 were from ATC and 8,820 were from other aircraft (pseudo pilots and the FAA B737-800 simulator, flown by “confederate” FAA pilots). The subject pilots were “talked to” approximately 1,380 times and “talked about” approximately 550 times.
Pilots erroneously responded on the radio eight (8) times during TPCS transmissions and internally discussed the transmission or queried one another, but did not transmit, four (4) times. Of those four (4), two (2) were TPCS and two (2) were TFID. This resulted in an approximate error rate, when “talked about” using TPCS formatting, of 1.45% (8 ÷ 550 = 0.0145). The erroneous response rate using the TFID formatting was zero (0). The four internally discussed, but not transmitted, events were noted above, but not included in the error rate, since no “error” was actually made.

7.6.1.2  Physiological Data Analysis

7.6.1.2.1  Eye Tracking

The 80ms dwell time is the minimum necessary for human perception to occur. Hence, the mapping filter is set to at least that time. The average saccade rate is 200ms. Typical mean fixation duration is between 180 to 300ms. The eye-tracking technology captured both eye (pupil) and head movement. The data mapping does not adequately display the distinction between head and individual pupil fixations/saccades. Each fixation is a snapshot in time. The software does capture the number of fixations/saccades (separately) and the duration of each.

Since post-data collection mapping is extremely time-consuming, analysts chose a representative sample of crew runs from which to perform semantic gaze mapping, see table 7-2. None of those runs were particularly singled out for mapping, but all the crews were represented. Sampling was conducted across low, medium, and high complexity for both the captain and first officer positions and for both the Airbus 330 and Boeing 737 subject populations, see images 7-2 to 7-9.

The images below are categorized as follows:

- Images 7-2 & 7-3 A330 captain pilot monitoring TFID;
- Images 7-4 & 7-5 A330 captain pilot flying TPCS with CDTI;
- Images 7-6 & 7-7 B737 captain pilot monitoring TFID with head-up display; and
- Images 7-8 & 7-9 B737 first officer pilot flying TPCS.

NOTE: Mapped fixations in those scenarios with an actively used head-up display are combined in the same key performance indicator blocks as forward field of view windshield.

Mapped pilot scan patterns are consistent with culturally established norms and we did not detect any anomalies that could be attributed directly to the nature of this test. The data indicates that the CDTI was referenced in a very limited scope and was associated exclusively with the TPCS format. Post data processing of eye-track data indicates that crews had primarily normal scan patterns. Our data confirms that pilots flying will dwell longer on the primary flight display while the pilot monitoring will dwell longer on the NAV display. Also, pilots monitoring have a more robust scan pattern while pilots flying tend to restrict their scan pattern to predominantly to the primary FOV, see images 7-2 to 7-9.

NOTE:Accepted commercial pilot duties universally require the pilot monitoring to monitor all radio traffic and respond as appropriate. Pilot flying duties require that crew member to primarily manage the control of the aircraft and secondarily monitor radio communication.
### Table 7-2: Eye-Track Mapping Schedule

<table>
<thead>
<tr>
<th>Complexity Levels by Scenario</th>
<th>Low - 1 and 4</th>
<th>Medium - 2 and 5</th>
<th>High - 3 and 6</th>
</tr>
</thead>
</table>

#### 21-28 July

<table>
<thead>
<tr>
<th></th>
<th>Low (EASY)</th>
<th>Medium</th>
<th>High (HARD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPCS (Callsign) CP</td>
<td>7/21-5-1</td>
<td>7/25 2-2</td>
<td>7/23 3-3</td>
</tr>
<tr>
<td>TFID (Letters) CP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPCS (Callsign) FO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFID (Letters) FO</td>
<td>7/24 6-4</td>
<td>7/25 5-5</td>
<td>7/22 6-6</td>
</tr>
</tbody>
</table>

#### 21-28 July

<table>
<thead>
<tr>
<th></th>
<th>Low (EASY)</th>
<th>Medium</th>
<th>High (HARD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPCS (Callsign) CP</td>
<td>7/21 6-4</td>
<td>7/22-1-5</td>
<td>7/23 4-6</td>
</tr>
<tr>
<td>TFID (Letters) CP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPCS (Callsign) FO</td>
<td>7/24-5-1</td>
<td>7/22-2-2</td>
<td>7/25 1-3</td>
</tr>
<tr>
<td>TFID (Letters) FO</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 29 Jul - 1 Aug

<table>
<thead>
<tr>
<th></th>
<th>Low (EASY)</th>
<th>Medium</th>
<th>High (HARD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPCS (Callsign) CP</td>
<td>8/1 6-1</td>
<td>7/30 3-2</td>
<td>8/1 5-3</td>
</tr>
<tr>
<td>TFID (Letters) CP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPCS (Callsign) FO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFID (Letters) FO</td>
<td>7/30 4-4</td>
<td>7/30 1-5</td>
<td>7/31 2-6</td>
</tr>
</tbody>
</table>

#### 29 Jul - 1 Aug

<table>
<thead>
<tr>
<th></th>
<th>Low (EASY)</th>
<th>Medium</th>
<th>High (HARD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPCS (Callsign) CP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFID (Letters) CP</td>
<td>7/30 4-4</td>
<td>8/1 1-5</td>
<td>7/30 2-6</td>
</tr>
<tr>
<td>TPCS (Callsign) FO</td>
<td>7/30 5-1</td>
<td>7/31 1-2</td>
<td>7/30 6-3</td>
</tr>
<tr>
<td>TFID (Letters) FO</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Image 7-4: Heat Map A330 Captain Pilot Flying Hard

Image 7-5: Areas of Interest A330 Captain Pilot Flying Hard
Image 7-8: Heat Map B737 First Officer Pilot Flying Hard

Image 7-9: Areas of Interest B737 First Officer Pilot Flying Hard
7.6.2 Qualitative Data Analysis

7.6.2.1 Subjective Questionnaire Responses

NOTE: Primary instruction to subject pilots as they respond to subjective questions; perspective should be a comparison against “normal” flying activities.

With the exception of the FAA stand-in pilot, questionnaire responses were successfully captured from all participants. Each pilot provided responses to all five questions after each run. A total of 102 responses to each question were captured. Not all pilots were familiar with or had experience with using the CDTI. For that reason, those specific crew members did not answer questions pertaining to the CDTI. Only five subject pilots were employed by an air carrier equipped with CDTI technology. For this reason only 30 responses were captured for this question.

The tables and figures associated with this portion of the report were organized in accordance with the following convention: Vertical versus horizontal bar alignment; ordinate scale; focused variables: mean values all; mean values by scenario complexity; mean values by pilot monitoring, and pilot flying.

Approach to Data Analysis:
Data analysis was viewed from an overall perspective which allowed confirmation of integrity and validity of responses. The mean of all responses fell in the center “typical” range for each question. Data was parsed further by aircraft type, pilot function (i.e. pilot flying or pilot monitoring), communication format (TPCS or TFID), scenario complexity, and combinations of these. The level of discrimination of the data was determined by two primary themes. First, the main data collection effort (DCE) objective that was stated in the test plan and second, by emergent aspects discovered from analysis during first notion analysis.

Figure Presentation Properties:
Several presentation properties remain consistent throughout the figures. First, consistency of color; colors used throughout the figures remain consistent within similar data categories. Blue is used to present TPCS data and red for TFID data. Second, the wall of the figure is color shaded to facilitate visual application of the vertical axis title (white on the bottom shading to green in the “typical” range shading to red at the top). Third, the vertical axis corresponds to the nine point scale with 1 on the bottom and 9 at the top with 5 representing center “typical”.

<<< Positive Typical Negative >>>

Fourth, questions are arrayed consistently in order from left to right on the horizontal axis.
The questions are presented here according to the left to right order used on the figures.

- **Difficulty** – Compared to your normal duties, rate the level of difficulty performing this arrival and approach;
- **Communication comfort** – Rate you comfort level with the communications procedures;
- **Ind workload** – Rate your perceived level of **individual workload** for this procedure from the standpoint of communication, coordination, and procedural habit patterns throughout the arrival procedure;
- **Crew workload** – Rate your perceived level of **crew workload** for this procedure from the standpoint of communication, coordination, and procedural habit patterns throughout the arrival procedure; and
- **CDTI comfort** – Rate your comfort level with this procedure utilizing the ADS-B/CDTI. (NOTE: this question does not appear on every figure; refer to specific figure explanations for analysis reasoning).
**Conceptual Set of the Questions:**
It will be useful to keep in mind the conceptual set of information directed for use by the subject pilot when answering the questions.

- **Question 1** directs the pilot to use a conceptual set of “normal duties”, which can include a broad number of aspects. As presented in figure 7-2, the mean is close to the median (i.e. a score of 5 on the figure indicating “typical”), indicating high uniformity of conceptual set across the pilot subjects when answering this question. While the question allows for many aspects to be included it does not necessarily prompt the subject to analytically think about what specifically is included in her set of “normal duties” while answering the question. One pilot’s “normal duties” may include much more, or less, than another’s. This is the nature of Human Factors subjective questions and is reason to limit inferential deduction to the threshold implicit in the question. In other words, since this question is broad in scope it is only appropriate to make equally broad inferences regarding the data. Choosing one aspect of “normal duties” and making an inference should be highly suspect as to do so would require reduction of concepts below that originally used for answering the question.

- **Question 2** specifically directed the subject pilot to use a conceptual set consisting of communication. As presented in figure 7-11 the mean is close to the median (a score of 5 on the figure indicating “typical”), indicating high uniformity of conceptual set across the pilot subjects when answering this question. Given the communication focus of this question, it is appropriate to evaluate potential communication correlations and inferences.

- **Question 3** directs the pilot to use a conceptual set of personal (individual) workload influencers to include communication, coordination, and procedural habit patterns. The mean for this question is close to the median (a score of 5 on the figure indicating “typical”) indicating high uniformity of conceptual set across subject pilots when answering this question. This question encompasses a broad range of aspects therefore inferential evaluations must also be commensurately broad.

- **Question 4** is similar to question 3 however question 4 asks the pilot to adjust the conceptual set to one of collective (crew) workload influencers. The mean for this question is close to the median (a score of 5 on the figure indicating “typical”) indicating high uniformity of conceptual set across subject pilots when answering this question. For similar reasons as those in question 3, inferential evaluations must commensurately broad with the question’s conceptual set.

- **Question 5** directs the pilot to use a conceptual set relationally tied to the ADS-B/CDTI. The mean for this question is close to the median (a score of 5 on the figure indicating “typical”) indicating high uniformity of conceptual set across subject pilots when answering this question. However, this question’s response (n-value) was only 30 and therefore limited its analysis to coarse parses. Therefore this question only appears in a few figures.
**Figures:**

**All Questions (Mean) (Figure 7-2):**

The figure presents the mean of the total responses to each question asked in the post run questionnaire. Mean datum for each question is near median, “typical”, indicating high validity of subject responses.

While overall a high validity of response is indicated it is noted that each question is slightly above the median line. Some reasonable influencing factors may be attributed to:

- Most pilots were not familiar with the specific arrival(s) used in the DCE;
- The nature of the scenario complexities were perceptively characterized as normal to above normal in individual responses and corroborated in post brief statements; and
- Responses to the question of communication comfort are closest to “5” which may indicate aspects other than processing unfamiliar communication phraseology were perceived as greater than typical; aspects such as other traffic density, convective weather impact, rerouting to join a different arrival, and downwind runway change. The dynamic nature of the these challenges, while not a-typical, added to the total secondary task activity which may have driven the perception of increased individual workload.

Caution is advised in attempting to extrapolate from data in this view as all aspects make up this graphical look, thus, it is impossible to supportively conclude any particular aspect’s individual contribution to the composite mean.
A330 vs B737 (Mean) (Figure 7-3):
The figure presents mean responses parsed by aircraft type. While all question's mean is near the median, the A330 means are slightly elevated over those of the B737. There are many factors contributing to this presentation however. One known contributing factor is that the A330 FMS algorithmic predictive altitude and speed control required above typical pilot intervention to meet designated arrival fix restrictions. According to pilot de-briefing statements, the constant pilot intervention to meet the arrival restrictions resulted in elevating the subject pilot assessment of work load.

This figure presents data for question 5, CDTI comfort. Notice that this question only applies to the A330 aircraft. Additionally, and not explicit in the figure is the limited n-value for this question. Question 5 n=30 while other question n-values are; A330 n=66, B737 n=36.
Figure 7-3: A330 vs B737 Mean
Pilot Flying vs Pilot Monitoring (Mean) (Figure 7-4):
The figure presents mean of each question parsed out by pilot function parse. The mean presents slightly above the median. The pilot flying primarily focuses on flying the aircraft and secondarily monitors the radios. Conversely the pilot monitoring primarily attends to radio communications and secondarily monitors flying of the aircraft. The pilot monitoring mean presents slightly elevated over the pilot flying mean. This may in part be influenced by the unfamiliarity of the NextGen communications in the Simulated Airspace System. Pilots during de-brief indicated that there was some minimally elevated cognitive processing associated with the unfamiliar communications phraseology. A deviation in the data is also present between the pilot monitoring and pilot flying. The scenarios involved a steady flow of activities throughout the arrival and approach in addition to the elements of NextGen communication phraseology. Much of the activity consistently involved tasks that fell to the pilot monitoring. The elevated pilot monitoring mean presented in this figure likely represents a combination of influencers, only some of which have been discussed here.

Figure 7-4: Pilot Flying vs Pilot Monitoring Mean
Pilot flying vs pilot monitoring, A330 vs B737 (Mean) (Figure 7-5):
The figure presents mean of responses according to a binary associated quartile parse. The data was parsed by aircraft type then further parsed by pilot flying and pilot monitoring functional duties. This figure presents a combination view of the “A330 vs B737 mean” and “pilot flying vs pilot monitoring mean” figure. From a statistical analysis perspective this parse limits the n-value represented by each bar (mean). For each A330 mean, n=33 and each B737 mean, n=18. The B737 n-values are considered to be at the lower margin for providing reliably useful information. For these reasons only broad general inferences should be made from this data view.

Note: The data presented for question 5 is identical (for each mean n=15) to that presented for this question in the “pilot flying vs pilot monitoring mean” figure.

![Figure 7-5: Pilot Flying vs Pilot Monitoring, A330 vs B737 Mean](image-url)
TPCS vs TFID (Mean) (Figure 7-6):
The figure presents mean of a TPCS vs. TFID parse. Within each question, TFID data is consistently lower than TPCS data. This TPCS/TFID relationship is corroborated by pilot de-briefing comments stating a preference for TFID usage when being referenced in ATC communications.

Note: Question 5, CDTI comfort represents a subset of the A330 pilot response mean.

Figure 7-6: TPCS vs TFID Mean
TPCS vs TFID (Mean) with Hi/Low (Figure 7-7):
The figure presents mean of a TPCS vs. TFID parse with highest and lowest data points. This is the same mean data (represented by the square data point markers) presented in “TPCS vs TFID mean” figure with the addition of hi/low bars which provide visual placement context of mean in the data range. Most of the means also coincide with the median of the hi/low points, indicating a balance of data throughout the data range. Regarding question 5, CDTI comfort, it is interesting to note the information this figure provides. The TPCS mean is in the lower quadrant between the hi/low points, indicating most responses fell below the median of the data yet two respondents registered “much more uncomfortable” (three standard deviations above “typical”). The TFID mean is coincident with the median of the data and presents within a narrow spread of responses; within one standard deviation above and below “typical”. The data presents a clear comfort preference for TFID when using the CDTI, which is corroborated with comments during de-brief sessions.

![Figure 7-7: TPCS vs TFID Mean with Hi/Low](image.png)
TPCS vs TFID, A330 vs B737 (Mean) (Figure 7-8):
The figure presents mean of responses according to a binary associated quartile parse. The data was parsed by aircraft type then further parsed by communication type. The data is presented in binary contrasting pairs of TPCS vs TFID within each aircraft type. This figure is a combination of the “A330 vs B737 mean” and “TPCS vs TFID mean” figures. Because of the low B737 data n-value (n=18 for each bar), it is considered to be at the lower margin for providing reliably useful information. Data for the A330 presents in accordance with a breakdown of other A330 parses, TPCS being elevated over TFID for each question. However, data for the B737 presents a new dynamic for questions 1 and 2. These two questions present a TPCS mean below TFID. This is a reversal of expectations which only presents in this parse combination. The parameters of this HITL DCE did not provide sufficient information to determine the reason for these mean presentations. Further, due to the small associated n-values, it is impossible to determine critical influencers without further testing in this area.

Note: The data presented for question 5 is identical (for each mean n=15) to that presented for this question in the “TPCS vs TFID mean” figure.

Figure 7-8: TPCS vs TFID, A330 vs B737
TPCS vs TFID, Pilot Flying vs Pilot Monitoring (Mean) (Figure 7-9):
The figure presents mean data of a binary associated quartile parse. The data is parsed by a four category matrix of pilot flying, pilot monitoring, TPCS, and TFID. Each parsing is then binarily associated by a TPCS vs TFID communication pair within a pilot flying vs pilot monitoring functional duty pair. This figure presents combined information of the “TPCS vs TFID” mean and the “pilot flying vs pilot monitoring mean” figures. TFID mean from the pilot flying consistently is elevated over TPCS while TFID mean from the pilot monitoring is consistently lower than TPCS. The pilot monitoring TFID mean is also consistently the lowest (most favorable) with in each set of four means.

Figure 7-9: TPCS vs TFID, Pilot Flying vs Pilot Monitoring Mean
Scenario Complexity (Mean) ((Figure 7-10):
The figure presents the spread of mean responses by scenario complexity. Mean datum presents around the mid-point “typical” indicating high confidence interval and validity of responses.

**Figure 7-10: Mean by Scenario Complexity**
Pilot Flying vs Pilot Monitoring—Mean by Scenario Complexity (Figure 7-11):
The figure presents mean according to pilot functional duties during high, medium, and low complexity scenarios. This parse was looked at because it separates pilot primary duties along communication functions. The pilot flying primarily focuses on flying the aircraft and secondarily monitors the radios. Conversely the pilot monitoring primarily attends to radio communications and secondarily monitors flying of the aircraft.

![Figure 7-11: Pilot Flying vs Pilot Monitoring – Mean Complexity](image-url)
TPCS vs TFID—Mean by Scenario Complexity (Figure 7-12):
The figure presents mean parse by scenario complexity within TPCS and TFID categories. The question addressing CDTI comfort is not included here because there is insufficient data at this level of parse to provide validity. Within each pair TFID presents favorable over TPCS.

7.6.2.2 Direct Observation
When given a first-party communication, directly to them (either directive or informational), subject pilots responded promptly in virtually all cases. This indicates a high level of vigilance and attentiveness to all communications. The inference would then be that they had the same level of attentiveness to third party communications, whether they verbalized or transmitted a response.

While test conditions were identical in both the A330 and B737 full flight simulators, observers noted that tailwinds seemed to cause A330 subject pilots to spend more time than B737 subject pilots monitoring and intervening in the FMS-managed descent profile, including significant use of speed brakes and selected speed, rather than managed speed.

7.6.3 Post Evaluation De-briefing Synopsis
During post-simulation de-briefings, 6 of 10 crews preferred the TFID format when being “talked about”, while the remaining four crews had no preference.
7.6.4 Major Theme Discussion (e.g. Strategies, Crew Interaction, Fatigue, Situational Awareness, etc.)

Workloads, based on the scenarios, appeared to be somewhat higher in the Airbus. This was primarily based on the increased need to intervene with SPEED BRAKES as well as SELECTED SPEED, in order to comply with all published constraints. Specifically, the Airbus tended to slow early in order to ensure compliance with the last published constraint on the STAR, which resulted in being too slow for an intermediate constraint. It should be noted that the Airbus FMS software version was not the latest version flown by the two air carrier pilot groups represented in the DCE. At least one air carrier technical pilot noted their more current FMS software would have made flying the STARs easier to accomplish.

Some pilots actively discussed the new phraseology format during some of the “talked about” events, both TPCS and TFID. For example, on numerous occasions, pilots would either visually or electronically (CDTI) reference the traffic being “talked to” when they were being “talked about.” Others said nothing during any of the events, but rarely or never missed a “talked to” radio call. Given that very few “talked to” transmissions were missed by the flight crews during the DCE, pilots appeared aware of being talked about, but that it simply was not a problem for them.

Use of DESCEND VIA phraseology was problematic. The Aeronautical Information Manual, para 5-4-1.a.2.(b) states… ‘Pilots cleared for vertical navigation using the phraseology “descend via” must inform ATC upon initial contact with a new frequency, of the altitude leaving, “descending via (procedure name),” the runway transition or landing direction if assigned, and any assigned restrictions not published on the procedure.’ Specifically, it appeared many pilots were unaware of the correct phraseology, or, as they noted in the debriefing, often fly into an airport that uses a variation on the phraseology, tailored for its airspace and operations. One pilot, who did not fly into the above-mentioned airport, correctly used the phraseology and noted during de-briefing his airline has been stressing the correct phraseology.

7.6.5 Subjective/Objective Conclusions

An area of particular interest is pilot communication vigilance, specifically as it pertains to third party radio communication. There was no direct way of determining if pilots were attending to and processing TPCS/TFID references. Crew vigilance was universally high, as evidenced by minimal observed errors in recognition and response to direct first-party calls. It should be kept in mind that “no reaction” to third party references does not necessarily imply a vigilance decrement, but perhaps points to a conscious effort not to respond. We investigated a potential connection between timely and accurate responses to first party calls and an inference to an equivalent level of vigilance to third party radio calls. Pilots gave timely and accurate responses to all direct first-party transmissions, pointing to a high degree of vigilance to radio traffic. A logical inference can be made that pilot vigilance to all third party calls was equally high.
8 Performance Data Analysis

From an observer perspective, considering that all flights were conducted almost entirely on autopilot, no significant aircraft performance implications were observed. Evaluation parameters include but are not limited to: autopilot disengagement; violation of published constraint airspeeds and/or altitudes; and execution of a missed approach/go around.

Note: One crew executed a go around due to potential violation of stabilized approach criteria and did not appear to be the result of being “talked about.” Several crews violated published constraint airspeeds in the A330. This is discussed in paragraph 7.6.4.1 above.
9 Conclusions

Overall, both controllers and pilots have indicated a slight preference of the TFID format over the TPCS. Regardless of that preference, data yielded from both the pilot and controller evaluations point that there is no confusion in the use of either phraseology formats nor is there an indication of a safety decrement in their use.
References

### Appendix A: Controller Test Anomaly Conditions and Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Level</th>
<th>B737</th>
<th>A330</th>
<th>Time</th>
<th>Incorrect</th>
<th>Not In Site</th>
<th>Blocked</th>
<th>Was That</th>
<th>Debate</th>
<th>Flip The</th>
<th>Adis Message</th>
<th>Other</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/22</td>
<td>1</td>
<td>Easy</td>
<td>SWA395</td>
<td>AWE216H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6:54</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11:56</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14:00</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B737 was blocked by A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>2</td>
<td>Hard</td>
<td>AAU1233</td>
<td>DAL153H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7:10</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9:30</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B737</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>3</td>
<td>Med</td>
<td>SWA539</td>
<td>FXX54MH</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3:50</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>4</td>
<td>Med</td>
<td>UAL1239</td>
<td>UPS5439H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3:50</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>5</td>
<td>Easy</td>
<td>DAL1591</td>
<td>DAL153H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A-1: Day 1A**

<table>
<thead>
<tr>
<th>Date</th>
<th>Level</th>
<th>B737</th>
<th>A330</th>
<th>Time</th>
<th>Incorrect</th>
<th>Not In Site</th>
<th>Blocked</th>
<th>Was That</th>
<th>Debate</th>
<th>Flip The</th>
<th>Adis Message</th>
<th>Other</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/22</td>
<td>1</td>
<td>Easy</td>
<td>SWA395</td>
<td>AWE216H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:24</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>2</td>
<td>Hard</td>
<td>AAU1233</td>
<td>DAL153H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9:30</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B737</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B737</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>3</td>
<td>Med</td>
<td>SWA539</td>
<td>FXX54MH</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3:50</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:30</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:30</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G/P</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/22</td>
<td>4</td>
<td>Med</td>
<td>UAL1239</td>
<td>UPS5439H</td>
<td>3:50</td>
<td>A300</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3:50</td>
<td>A300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3:50</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3:50</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10:30</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11:00</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12:30</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:00</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:00</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:00</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:00</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13:00</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A-2: Day 1B**
### Figure A-3: Day 2A

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Level</th>
<th>Flight</th>
<th>Time</th>
<th>Incorrect Video Feedback</th>
<th>Not In Site</th>
<th>Blocked Voice XMT</th>
<th>Was That Call For Me?</th>
<th>Other</th>
<th>Flight The Switch</th>
<th>ADS Message</th>
<th>Other</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/23 1 1</td>
<td>Easy</td>
<td>SWA365</td>
<td>15:00</td>
<td>B737</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ATC did not catch</td>
</tr>
<tr>
<td>4/23 1 2</td>
<td>Hard</td>
<td>AAL1239</td>
<td>16:30</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>No time - On downwind</td>
</tr>
<tr>
<td>4/23 1 3</td>
<td>Med</td>
<td>SWA329</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>Corrected by ATC</td>
</tr>
<tr>
<td>4/23 1 4</td>
<td>Med</td>
<td>UAL1229</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>G/P Incorrect Call Sign used on Polish</td>
</tr>
<tr>
<td>4/23 1 5</td>
<td>Hard</td>
<td>DAL1681</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>ATC did not catch mistake</td>
</tr>
<tr>
<td>4/23 1 6</td>
<td>Easy</td>
<td>UAL677</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>ATC responded with instructions to switch freq</td>
</tr>
</tbody>
</table>

### Figure A-4: Day 2B

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Level</th>
<th>Flight</th>
<th>Time</th>
<th>Incorrect Video Feedback</th>
<th>Not In Site</th>
<th>Blocked Voice XMT</th>
<th>Was That Call For Me?</th>
<th>Other</th>
<th>Flight The Switch</th>
<th>ADS Message</th>
<th>Other</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/23 2 1</td>
<td>Easy</td>
<td>SWA365</td>
<td>15:00</td>
<td>B737</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1st party</td>
</tr>
<tr>
<td>4/23 2 2</td>
<td>Hard</td>
<td>AAL1239</td>
<td>15:00</td>
<td>B737</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ATC did not catch</td>
</tr>
<tr>
<td>4/23 2 3</td>
<td>Med</td>
<td>SWA329</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>ATC responded</td>
</tr>
<tr>
<td>4/23 2 4</td>
<td>Med</td>
<td>UAL1229</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>Ignored by ATC</td>
</tr>
<tr>
<td>4/23 2 5</td>
<td>Hard</td>
<td>DAL1681</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>ATC responded</td>
</tr>
<tr>
<td>4/23 2 6</td>
<td>Easy</td>
<td>UAL677</td>
<td>15:00</td>
<td>A330</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A330</td>
<td></td>
<td>ATC called the B737 &quot;Heavy&quot;</td>
</tr>
</tbody>
</table>

TFID Issued on [August 2015] [Flight Operations Simulation Branch, AFS-440]
<table>
<thead>
<tr>
<th>Date/1</th>
<th>Level</th>
<th>Time</th>
<th>Incorrect</th>
<th>Not In Site</th>
<th>Blocked</th>
<th>Was That For Me?</th>
<th>Debate Instructions</th>
<th>Failed/Success</th>
<th>Other</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/24 1</td>
<td>Easy</td>
<td>9:15</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 1</td>
<td>Hard</td>
<td>16:44</td>
<td>G/P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>29:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 2</td>
<td>Hard</td>
<td>1:55</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4:25</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6:40</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17:54</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 3</td>
<td>Med</td>
<td>5:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5:30</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7:15</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13:30</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17:00</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 4</td>
<td>Med</td>
<td>5:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5:15</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6:30</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13:30</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>21:30</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 5</td>
<td>Hard</td>
<td>2:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5:50</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6:00</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15:10</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17:30</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 6</td>
<td>Easy</td>
<td>5:00</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9:30</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A-5: Day 3A**

<table>
<thead>
<tr>
<th>Date/2</th>
<th>Level</th>
<th>Time</th>
<th>Incorrect</th>
<th>Not In Site</th>
<th>Blocked</th>
<th>Was That For Me?</th>
<th>Debate Instructions</th>
<th>Failed/Success</th>
<th>Other</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/24 1</td>
<td>Easy</td>
<td>7:48</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19:40</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>21:00</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 2</td>
<td>Hard</td>
<td>11:52</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9:05</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13:13</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 3</td>
<td>Med</td>
<td>4:15</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8:50</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15:41</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 4</td>
<td>Med</td>
<td>3:00</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3:15</td>
<td>GP/B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8:35</td>
<td>GP/B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15:35</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 5</td>
<td>Hard</td>
<td>2:50</td>
<td>B737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4:30</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9:30</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9:45</td>
<td>GP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16:52</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/24 6</td>
<td>Easy</td>
<td>7:40</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16:56</td>
<td>A330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A-6: Day 3B**
Appendix B: Pilot Scenarios

Figure B-1: Scenario 1 – Low Complexity

<table>
<thead>
<tr>
<th>Talked To</th>
<th>Talked About</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>CAVS Clearance</td>
<td>CAVS Clearance</td>
</tr>
</tbody>
</table>

Table B-1: Scenario 1 – Easy
Imagine B-1: Approach Chart – Easy Scenario
Figure B-2: Scenario 2 – Medium Complexity

<table>
<thead>
<tr>
<th>Talked To</th>
<th>Talked About</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>CAVS Clearance</td>
<td>CAVS Clearance</td>
</tr>
</tbody>
</table>

Table B-2: Scenario 2 Medium

- Night
- VMC/Light Chop (40%)
- Light radio traffic
- Over FRNCH
- FL250
- 280 KIAS
- HDG 094
- Medium Weight
- Vector off arrival

- 4-7 Traffic Call Outs
- 1 CAVS Approach Clearance
- Moderate opposite (talk to/about) communications
ARRIVAL ROUTE DESCRIPTION

From FRNCH on track 094° to cross SKARF at or above FL220 and at or below FL250 at 270K, thence as depicted to WASKL, then on track 173°. Expect RADAR vectors to final approach course.

LOST COMMUNICATIONS: In the event of lost communications prior to runway assignment: Aircraft landing DEN execute the ILS RWY 35L, aircraft landing BKF execute the ILS RWY 32, aircraft landing BJC execute the ILS RWY 35R.

NOTE: Chart not to scale.

Image B-2: Approach Chart – Medium Scenario
Figure B-3: Scenario 3 – High Complexity

<table>
<thead>
<tr>
<th>Talked To</th>
<th>Talked About</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Traffic call out</td>
<td>Traffic call out</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>Talked about</td>
<td>Talked to</td>
</tr>
<tr>
<td>CAVS Clearance</td>
<td>CAVS Clearance</td>
</tr>
</tbody>
</table>

Table B-3: Scenario 3 Hard
Image B-3: Approach Chart – Hard Scenario

NOTE: RADAR required.
NOTE: RNAV 1.
NOTE: DME/DME/IRU or GPS required.
NOTE: Turboprops only.
NOTE: Expect runway assignment on initial contact with Denver TRACON.
NOTE: The corresponding RNAV STAR is KAIE.
   Expect KAIE when Denver is landing south.
NOTE: Aircraft landing BKF/BJC, expect RADAR vectors to final approach course.

ARRIVAL ROUTE DESCRIPTION
From FRNCH on track 094° to cross SKARF at or above FL220 and at or below FL250 at 270K, thence as depicted to WASKL, then on track 173°. Expect RADAR vectors to final approach course.

LOST COMMUNICATIONS: In the event of lost communications prior to runway assignment: Aircraft landing DEN execute the ILS RWY 35L, aircraft landing BKF execute the ILS RWY 32, aircraft landing BJC execute the ILS RWY 35R.

TFID Issued on [August 2015] Page 71 of 79
[Flight Operations Simulation Branch, AFS-440]
Appendix C: Controller Demographic Questionnaire

Date: _______________  Controller # ___________

1. What facility do you currently perform as a controller? ________________
2. Number of years as a certified controller? ________________
3. Are you currently using Fusion? ________________
4. Are you familiar with ADS-B? ________________
5. Approximately how many years have you used the ADS-B technology? ________________
6. How comfortable are you in the use of the ADS-B?

________________________________________________________________________
________________________________________________________________________
Appendix D: Pilot Demographic Questionnaire
DATE: __________________ CREW #:________________ CP/FO

1. What airline are you currently employed by?

2. Are you current and qualified (have landing currency) in the aircraft you flew in the simulation?

3. Are you currently flying line operations for your company?

4. If you are not current and qualified in the aircraft of question 2, what aircraft are you currently flying for your employer? Approximately how many hours do you have in the aircraft in which you are currently qualified?

5. Are you familiar with ADS-B technology?

6. Have you used EFB/CDTI Technology on the line?
Appendix E: Controller Post-Run Questionnaire

1. Compared to your normal duties, rate the level of difficulty performing controller duties with the required change in call sign usage and phraseology.

<table>
<thead>
<tr>
<th>Much Easier</th>
<th>Same as Typical Operation</th>
<th>Somewhat More Difficult</th>
<th>Much More Difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

2. Rate your level of comfort using the call sign format in this session.

<table>
<thead>
<tr>
<th>Very Comfortable</th>
<th>Somewhat Comfortable</th>
<th>Same as Typical Operation</th>
<th>Somewhat Uncomfortable</th>
<th>Very Uncomfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

3. Rate how timely you felt your instruction were, given the call sign format and phraseology you were instructed to use.

<table>
<thead>
<tr>
<th>Faster</th>
<th>Same as Typical Operation</th>
<th>Slower</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

4. Rate your perceived level of individual workload, from the standpoint of mental demand (e.g. looking, searching, thinking, deciding, communicating etc.) for this operation.

<table>
<thead>
<tr>
<th>Lower</th>
<th>Same as Typical Operation</th>
<th>Higher</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

5. As compared to other parallel and/or simultaneous instrument approaches, rate the collective workload (final monitor, tower and coordinator interaction, etc.) for this operation.

<table>
<thead>
<tr>
<th>Lower</th>
<th>Same as Typical Operation</th>
<th>Higher</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Appendix F: Pilot Post-Run Questionnaire

1. Compared to your normal duties, rate the level of difficulty performing pilot duties with the required change in call sign usage and phraseology.

<table>
<thead>
<tr>
<th>Much Easier</th>
<th>Somewhat Easier</th>
<th>Same as Typical Operation</th>
<th>Somewhat More Difficult</th>
<th>Much More Difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

2. Rate your comfort level with this procedure and the use of the ADS-B/CDTI.

<table>
<thead>
<tr>
<th>Very Comfortable</th>
<th>Somewhat Comfortable</th>
<th>Same as Typical Operation</th>
<th>Somewhat Uncomfortable</th>
<th>Very Uncomfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

3. Rate your comfort level with the communications procedures.

<table>
<thead>
<tr>
<th>Very Comfortable</th>
<th>Somewhat Comfortable</th>
<th>Same as Typical Operation</th>
<th>Somewhat Uncomfortable</th>
<th>Very Uncomfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

4. Rate your perceived level of individual workload for this procedure from the standpoint of communication, coordination, and procedural habit patterns throughout the arrival procedure.

<table>
<thead>
<tr>
<th>Much Lower Workload</th>
<th>Somewhat Lower Workload</th>
<th>Same as Typical Operation</th>
<th>Somewhat Higher Workload</th>
<th>Much Higher Workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

5. Rate your perceived level of crew workload for this procedure from the standpoint of communication, coordination, and procedural habit patterns throughout the arrival procedure.

<table>
<thead>
<tr>
<th>Much Lower Workload</th>
<th>Somewhat Lower Workload</th>
<th>Same as Typical Operation</th>
<th>Somewhat Higher Workload</th>
<th>Much Higher Workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Appendix G: Post-Evaluation De-briefing – Controllers

DATE: _________       Controller Pair :______

1. Were you comfortable with the required callsign phraseology in the evaluation?
   _______ Why?/Why Not? _______________________________________________
   ___________________________________________________________________

2. Do you feel this phraseology impeded your ability to effectively communicate with
   aircraft in an efficient and timely manner? Explain
   ___________________________________________________________________
   ___________________________________________________________________

3. What additional mental or physical requirements, if any, were imposed on you during
   this simulation?
   ___________________________________________________________________
   ___________________________________________________________________

4. Which part of the simulation was more difficult (in trail spacing, traffic density,
   course deviation, and coordination/communication)? Why?
   ___________________________________________________________________
   ___________________________________________________________________
   ___________________________________________________________________

5. Is it more important to ensure the specified phraseology as directed by this test is
   used, or is it more important to transmit a call as soon as possible regardless of the
   wording?
   ___________________________________________________________________
   ___________________________________________________________________
   ___________________________________________________________________

6. Which callsign phraseology do you feel is optimal?    Which do you prefer?
   ___________________________________________________________________
   ___________________________________________________________________

7. With respect to this operation, do you have any suggestions for the following:
   a. Controller training?
      ___________________________________________________________________
   b. Areas that require further evaluation?
      ___________________________________________________________________
8. Rate the realism of this system’s (if unrealistic say why, e.g., inconsistent, jerky, etc.):
   
a. Video Display
   _________________________________________________________________

b. Display of track
   movement________________________________________________________

c. Audio__________________________________________________________

d. What would make the system more realistic?
   ______________________________

9. Do you have any other comments about anything you observed during the
   simulations?
   _________________________________________________________________
   _________________________________________________________________
   _________________________________________________________________

10. Do you have any suggestions for the use of either phraseology option in the future?
    Training? Equipment?
    ________________________________________________________________
    ________________________________________________________________
Appendix H: Post Evaluation De-briefing – Pilots
DATE: _________  CREW #___________  OBSERVER: ______________________

1. Were there significant differences between this simulator compared to the aircraft you fly for your company and did it impact your performance?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

2. If applicable, was the position/orientation of the Electronic Flight Bag/CDTI in our simulator consistent with what you are accustomed to/comfortable to flying with?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

3. Did the EFB position effect your performance in the simulation? Explain
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

4. Overall, did you feel comfortable with this operation? ______ Why or Why Not?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

5. What additional mental or physical requirements were imposed on you during this operation? Were there any changes to your workload (mental or physical)?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

6. In actual operations, how comfortable would you be in performing this operation in less than optimal conditions (e.g. Poor weather, no A/P or A/T, crosswinds, etc.)?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

7. Do you think that the phraseology used in this test would be acceptable during actual operations? _________ If not, please explain
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
8. Did you prefer either of the phraseology options over the other?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

9. Did either of the phraseology options affect your performance or create confusion?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

10. Did you have any trouble understanding the intent of the controller’s instructions? Who he/she was referring to?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

11. Do you have any suggestions for the use of either phraseology option in the future? Training? _____________________________
Equipment? ____________________________________________