

Use of Data Comm by Flight Crew in High-Density Terminal Areas

Brian T. Baxley¹

NASA Langley Research Center, Hampton, VA 23681

Robert M. Norman, Ph.D.²

Boeing Research and Technology, Hampton, VA 23681

Kyle K. E. Ellis³

University of Iowa and NASA Graduate Student Researchers Program

Kara A. Latorella, Ph.D.,⁴ James R. Comstock, Ph.D.,⁵ and Cathy A. Adams⁶

NASA Langley Research Center, Hampton, VA 23681

This paper describes a collaborative FAA and NASA experiment using 22 commercial airline pilots to determine the effect of using Datalink Communication (Data Comm) to issue messages in busy, terminal area operations. Four conditions were defined that span current day to future flight deck equipage levels (voice communication only, Data Comm only, Data Comm with Moving Map Display, Data Comm with Moving Map displaying taxi route), and each condition was used to create an arrival and a departure scenario at the Boston Logan Airport. These eight scenarios were repeated twice for a total of 16 scenarios for each of the eleven crews. Quantitative data was collected on subject reaction time and eye tracking information. Questionnaires collected subjective feedback on workload and acceptability to the flight crew for using Data Comm in a busy terminal area. 95% of the Data Comm messages were responded to by the flight crew within one minute; however, post experiment debrief comments revealed almost unanimous consensus that two minutes was a reasonable expectation for crew response. Eye tracking data indicated an insignificant decrease in head-up time for the Pilot Flying when Data Comm was introduced; however, the Pilot Monitoring had significantly less head-up time. Data Comm workload was rated as operationally acceptable by both crew members in all conditions in flight at any altitude above the Final Approach Fix in terms of response time and workload. Results also indicate the use of Data Comm during surface operations was acceptable, the exception being the simultaneous use of voice, Data Comm, and audio chime required for an aircraft to cross an active runway. Many crews reported they believed Data Comm messages would be acceptable after the Final Approach Fix or to cross a runway if the message was not accompanied by a chime and there was not a requirement to immediately respond to the uplink message.

¹ Research Engineer, Crew Systems and Aviation Operations Branch (CSAOB), 24 W Taylor St, Hampton, VA 23681, AIAA Senior Member

² Technical Fellow, CSAOB, 24 W Taylor St, Hampton, VA 23681

³ Graduate Research Engineer, CSAOB, 24 W Taylor St, Hampton, VA 23681, AIAA Student Member

⁴ Research Engineer, CSAOB, 24 W Taylor St, Hampton, VA 23681

⁵ Research Engineer, CSAOB, 24 W Taylor St, Hampton, VA 23681

⁶ Research Engineer, CSAOB, 24 W Taylor St, Hampton, VA 23681

Nomenclature

<i>CDU</i>	=	Control Display Unit
<i>CPDLC</i>	=	Controller Pilot Datalink Communications
<i>D-Taxi</i>	=	Datalink Taxi
<i>HITL</i>	=	Human In The Loop
<i>IFD</i>	=	Integration Flight Deck
<i>KBOS</i>	=	Boston Logan International Airport
<i>MMD</i>	=	Moving Map Display
<i>ND</i>	=	Navigation Display
<i>PF</i>	=	Pilot Flying
<i>PM</i>	=	Pilot Monitoring
<i>SA</i>	=	Situation Awareness

I. Introduction

THE Next Generation Air Transportation System (NextGen) Concept of Operations supports employment of an electronic Data Communications (Data Comm) system to effect communication of non-time critical information and events between flight deck and ground facilities. Transferred information could include expected taxi and flight clearances, weather, and real-time trajectory control. Deploying such a system is intended to sustain the expected increase in air traffic by 2025 and to assist Air Navigation Service Providers in managing air traffic more efficiently.¹ For flight crews, Data Comm serves as a means to share weather, airport surface operations, and trajectory-based operations information. Additionally, Data Comm and associated automation on the flight deck allows real-time data to be transmitted, thereby improving decision making for controller and pilots.¹ The FAA plans on implementing this Data Comm system in three segments. Segment One is scheduled from the years 2012 – 2016 and will be characterized by current equipment and the use of voice communications as the primary mode. Segment Two is scheduled in the years 2017 – 2022 and will be characterized by the gradual transition from voice to Data Comm (both in terms of procedures and equipment) during the segment time period. Segment Three is scheduled to begin in 2023 and is characterized by Data Comm being the primary mode of communication.

The Federal Aviation Administration (FAA) and National Aeronautics and Space Administration (NASA) partnered to conduct a Human-in-the-Loop (HITL) simulation to research Controller Pilot Datalink Communication (CPDLC), a subset of the Data Comm functionality, to understand how it will contribute to the ultimate NextGen goal of increased efficiency and capacity in the Segment Two timeframe. This paper describes the design and results from the NASA Langley Research Center experiment conducted in February and March 2010. The experiment used the high-fidelity Integration Flight Deck (IFD) simulator with 22 current, commercial airline pilots flying a complex and operationally realistic simulation of arrivals and departures from Boston Logan Airport in a heavy traffic density environment.

II. The Promise and Challenge of Data Comm to Flight Crew

A comprehensive review of Data Comm research in 1999 identified several issues with voice communication: sequential data transmission, background noise and dialect, frequency congestion, and long or complex messages being prone to being misunderstood. This study also identified benefits of Data Comm communication: higher efficiency, unloading of flight crew memory, improved message delivery time, improved transfer of information to other Air Traffic Control (ATC) and flight deck systems. This study also listed challenges of Data Comm that include: reduced situation awareness (SA) due to loss of “party-line” voice communication, increased head-down time, inability to multi-task while responding to Data Comm.² However, research by the FAA indicates the use of graphical displays may aid in increasing SA and reducing head-down time.³

Three recent and significant operational validations were conducted in Europe within the past ten years. The LINK2000+ research established guidelines based on HITL simulations using controllers and pilots in European enroute airspace, in particular the requirement that voice be used for all time-critical and safety-related communication. Other findings include delay in communication response, lack of flexibility in composing Data Comm messages, and loss of SA for party-line communication.⁴ Over 900 revenue flights at the Brussels airport from August 2006 through February 2007 participated in operational trials (the so-called D-TAXI tests) exploring procedures to improve productivity and safety while using non-time critical messages for a medium to high taxi path

complexity. Push-back, start-up, and taxi Controller-Pilot Datalink Communication (CPDLC) messages were sent by ATC and responded to by the crew using CPDLC, however for operational and safety reasons the crews also responded using voice. Overall the pilot acceptance, based on debrief comments, was high and continued to increase as the experiment continued. However, response time was longer although it was stated it was not operationally significant. Results from pilot debrief found the requirement for the flight crew to respond with both Data Comm and voice was considered impractical.⁵ Researchers from 24 multi-national partners conducted EMMA (2004-2006) and EMMA2 (2006-2009). Multiple simulation platforms and operational test locations were used to explore the holistic, integrated air-ground system. For the taxi tests in Prague, Milan, and Toulouse, the flight crews had a Moving Map Display (MMD) as well as surface alerts for other traffic and runway incursions, and CPDLC was used to transmit taxi instructions. Safety analysis conducted by the consortium resulted in the requirement that voice communication always took precedence over Data Comm. EMMA results indicated taxi time and voice communication was reduced, while EMMA2 concluded that CPDLC for taxi operations under these conditions was technically and operationally feasible.⁶⁻⁹

Significant research has been conducted in Data Comm, but emphasis in emerging NextGen operational conditions is needed. For instance, NASA flight tests at Denver's Stapleton Airport concluded that Data Comm was more accurate than voice and lowered pilot workload. Cockpit equipment included the ability to automatically load the ATC instruction from the Control Display Unit (CDU) into the Flight Management System (FMS).¹⁰ Another flight test in 1991 by NASA Langley with a Boeing 737 and seven crews showed a reduction in confusion, errors, and need for message repetition when the ability to "auto-load" the ATC instruction into the FMS was available.^{11, 12} A 2002 NASA simulation study found messages sent via Data Comm reduced time spent writing clearances and improved the crews' ability to understand the message on the first attempt.¹³ Another NASA study indicated that an electronic moving map significantly enhanced SA when using Data Comm, and reduced both intra-cockpit and controller-pilot voice communications.¹⁴ The present work does not, however, consider the issues and challenges associated with trajectory negotiation within NextGen.^{15,16}

In summary, voice communication should be used for time-critical, safety related, and non-normal situations since generally, the communication occurs more quickly, and voice is more flexible than Data Comm in transmitting non-typical information. Voice communication has usually been given priority over Data Comm to ensure there is no ambiguity between the two types of communication, and the use of Data Comm reduces the need for communication via voice between controllers and pilots. The acceptability of Data Comm by the flight crew was heavily influenced by the installation location and capability of the Data Comm equipment. Data Comm messages are less prone to being misunderstood, especially in voice communications impacted by accents, dialects, and language/cultural influences and biases, missed communication or inattention, and the memory burden required for voice. Data Comm messages require longer flight crew response times, in part due to interpreting the message prior to responding and the need for one crew member to brief the other, and require more head-down time for the flight crew. Displaying Data Comm messages graphically reduced workload and potentially offsets the loss of SA normally acquired through listening to voice communication of surrounding traffic.

III. Experiment Design Overview

The FAA/NASA Interagency Agreement for this experiment specified that the focus of research was CPDLC uplink messages (ATC to flight crew) and, in particular, if pilots find the use of Data Comm in a busy terminal area to be acceptable, and should controllers be inhibited from sending taxi or expected taxi clearances due to high flight crew workload. Due to the significantly different flight deck roles and responsibilities, separate data were collected and analysis metrics applied for the Pilot Flying (PF) and Pilot Monitoring (PM).

A. Experiment Hypotheses

Three hypotheses were identified:

- H1: Pilot workload and SA will differ significantly between communication modalities.
- H2: Pilot workload and SA will differ significantly between display modes when using Data Comm.
- H3: Pilots will rate the Data Comm used within this experiment as operationally acceptable.

In addition to addressing these hypotheses, the design of the study also examined:

- Message response times by type of CPDLC Data Comm message
- Objective data and subjective responses broken down by inflight and surface segments
- Objective data and subjective responses broken down by arrival and departure scenarios

- Workload and SA by crew member (PF and PM)
- Acceptability of Data Comm messages at "High", "Medium", and "Low" altitude bands during arrivals
- Assessments of head-up time for each crew member across the experimental conditions

B. Experiment Design

The experiment design was created from two communication modalities (voice vs. Data Comm) with three flight deck display concepts (paper charts, MMD, and MMD with route). As shown in Table 1, only four cells were evaluated. Unpopulated cells were removed from the experiment due to available time, funding, and not being essential for the primary purpose of the experiment.

The column "Data Comm + Exception" refers to where the primary communications modality is Data Comm, but voice is used for initiation of aircraft movement, aircraft check-in on Tower and Ground frequencies, runway crossing and clearance for takeoff or landing. The populated cell for "Paper Airport Diagram" and "Voice" was intended to be a baseline case, in terms of workload and situation awareness, representing typical airline transport operations in a present-day (2010) environment. Thus, incorporation of expected or amended taxi clearances in a current day voice communications environment, while not unheard of, represents atypical operations and therefore was not implemented in this experiment design. For Data Comm, the scenarios were designed to be extremely intrusive, with CPDLC messages sent within two minutes, and given in conjunction with out the window traffic (arrival scenarios) or with aircraft weight change and takeoff speed recalculations (departure scenarios). The Data Comm + Exception configurations held constant the modality of messages (Data Comm) but varied the display methodology over three display conditions: paper, moving map, and moving map with route on the navigation display.

Eight scenarios were developed for each subject crew using each of the populated cells (Table 1) and using one of two departures and one of two arrival tasks.

Table 1. Experimental Design Matrix

	Voice	Data Comm + Exception
Paper Airport Diagram	✓	✓
MMD Only		✓
MMD + Route		✓

These eight scenarios were flown twice, so that each crew flew 16 runs. Runs were ordered so that modality/display methodology conditions were not repeated within 3 runs, with arrivals and departures alternating as much as possible. Run conditions for the first eight runs were replicated exactly for the second set of eight runs for all crews. However, different run orders were assigned to different crews in order to counterbalance the serial position of scenarios over the course of the experiment.

C. Variables and Metrics

The first Independent Variable assessed the differences in pilot acceptability of communications using two different modalities (Voice and Data Comm), and the second investigated the effect of map display methodology on the acceptability of Data Comm. The altitude at which four of the CPDLC messages were given during the arrival scenarios was also varied (from 17,000' MSL to 5,000' MSL).

The Dependent Variables were selected to evaluate the acceptability to the flight crew of Data Comm use in the terminal maneuvering area during arrival and departure (surface) operations. A subset of the measures used in the experiment include:

- Data Comm message response times
- Head-up time of the PF and PM
- Bedford Workload Scale¹⁷
- Acceptability (subjective, self-rating)

D. Scenario Environment

The arrival scenarios were based on the NORWICH THREE (southern arrival) and SCUPP FOUR (eastern arrival) (see Figure 1) into Boston-Logan International Airport (FAA identifier KBOS). All arrivals were hand-flown, with auto-throttles engaged to increase pilot workload during the test. Scenarios, from parking gates, taxi

routing, and arrival routing, was chosen to match those used by the FAA William J. Hughes Technical Center for a corresponding controller HITL experiment to the extent possible and practical.

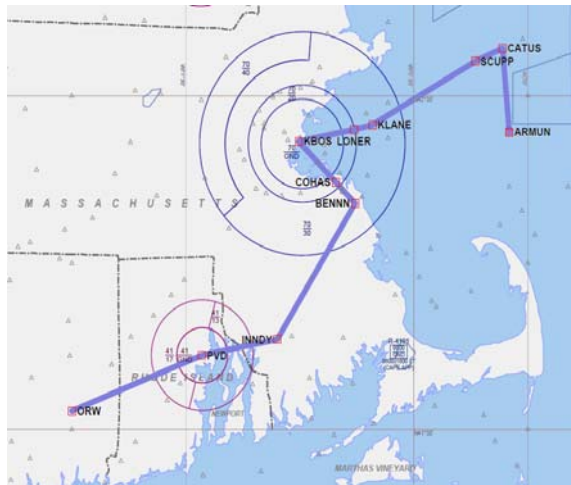


Figure 1. Arrival procedures and airspace
communication from other aircraft were pre-recorded, requiring approximately 170 dynamic aircraft and 1240 audio files for the 18 scenarios.

E. Research Facilities

The IFD full-mission simulator shown in Figure 2 replicates a Boeing 757-200 aircraft. The cockpit includes standard ship's instruments representative of a line operations Boeing 757-200 aircraft. The Main Instrument Panel contains Primary Flight Displays (PFD) and Navigation Displays (ND), Engine Indication and Caution Alerting System, flight instruments (airspeed, altitude, attitude, etc), as well as standby altimeter and gear lever. The Center Control Stand consists of B-757 throttle quadrant, flap and speed brake controls, reverse thrust, spoiler handles, dual FMS CDUs, and several electronic panels for controlling the PFD and ND, as well as researcher-specified systems. The IFD houses a standard Mode Control Panel under the glare shield, and a complete Overhead Panel.

The moving map display concepts, when flown, were shown on the each pilot's ND. The ND was a color Cathode Ray Tube device subtending approximately 5.75 inches vertically by 4.75 inches horizontally.

The cockpit's visual system is a panorama system using five 1440 x 1024 pixel resolution video projectors that provide 200° horizontal by 40° vertical field-of-view. A Rockwell Collins EP-1000 Boston-Logan database drove the out-the-window projection of the airport surface, taxiways, runways, buildings, obstructions, signs, and airport terrain and cultural features in a day, visual meteorological conditions setting. Up to 20 aircraft were shown in the arrival and surface taxi scenarios, and this traffic was accurately projected in the out-the-window displays, and shown on the moving map display if appropriate for that run condition.



Figure 2. Integration Flight Deck (IFD)

A ten-camera remote oculometer system was installed to support unobtrusive collection of eye tracking and head position data for both flight crew subjects. The cameras use facial features detection to track the subject's head with six degrees of freedom and corneal reflection to calculate left and right eye gaze vectors.

F. Data Comm Messages and Graphical Display

Data Comm messages were formatted in accordance with RTCA DO-305, DO-219, and DO-269, and based on a Boeing 747 FANS 1/A implementation. Each crew received 96 Data Comm messages during their 16 scenarios, with the data analysis focused primarily on the 54 messages of three specific types:

- taxi clearances (from terminal to departure runway, and from landing roll-out to the terminal area)
- amended taxi clearances (change to current taxi route, and to cross an active run during departure scenario)

- expected taxi clearances (given by ATC for flight crew planning purposes)

For realism and to preclude the subject pilots from always expecting these messages, other CPDLC uplink messages were interspersed in the scenarios, such as changes to altimeter settings, notification of a new ATIS recording, and issuing a frequency change to the next controller. In addition to responding to all uplink messages appropriately (ROGER, WILCO, UNABLE, or STANDBY), the subject pilots were required to use CPDLC downlink messages such as requesting pushback, requesting engine start, and requesting their taxi clearance.

The altitude at which these messages were given during arrival scenarios was varied by altitude. Two “Expected Taxi” clearances, an ATIS update, and an altimeter change were given between 16,000’ and 14,000’ MSL to test acceptability at the high altitude bands, between 10,000’ and 8000’ MSL for medium altitude testing, and 7000’ to 5000’ MSL for low altitude testing. The frequency change to tower was always given at approximately 4000’ MSL.

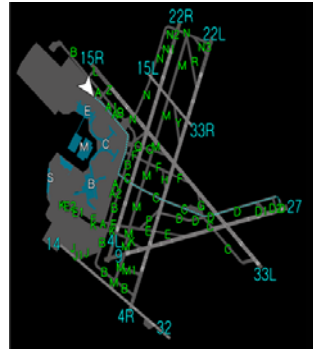


Figure 3. Expected Taxi Message and Display

In Figure 3, an example of a Data Comm “Expected Taxi” message on the CDU is shown for a departure scenario, and the corresponding MMD on the ND. When the CPDLC uplink is received, a chime sounds in the cockpit and the PM (always the First Officer in this experiment) depresses the ATC button on the CDU to visually read the message before briefing it to the PF. On Page One of the message, the PM loads the text clearance onto the MMD as a graphical route (if the scenario is the condition that includes displaying the route graphically). The “Expected Taxi”

message is displayed as a dotted cyan line from the terminal gate to the departure runway, with no red hold short lines for any runway (to visually reinforce that an “Expected Taxi” message is not a taxi clearance, and is for planning purposes only). The same procedure applies for an arrival scenario where ATC sends an uplink message of the taxi route from runway to terminal for planning purposes. The PM responds to ATC with a ROGER for informational-type CPDLC messages such as expected taxi clearances, changes to the ATIS or altimeter settings.

The “Taxi” and “Amended Taxi” clearances are accessed in a similar fashion by the PM. A chime sounds when the uplink is received, the PM reads and briefs the message, then loads the graphical representation on the MMD (if appropriate) from Page One (Figure 4). The taxi clearance is shown as a dotted white line from ownship position to the first active runway, where it ends in a solid red bar representing the Hold Short line (not clearly visible in these pictures at the intersection of Taxiway D and Runway 33L, however very visible in the IFD during the experiment). The rest of the route past the red Hold Short bar remains dotted cyan. When the PM acknowledges the taxi clearance with a WILCO on Page Two of the message, the dotted white line turns to a solid magenta.

In the voice condition used to simulate operational procedures in effect today, the subject flight crew were cleared to begin taxing the aircraft once they acknowledged and read back the ATC taxi clearance. However the Data Comm environment is slightly different. A Data Comm taxi clearance is just the route, and not permission to

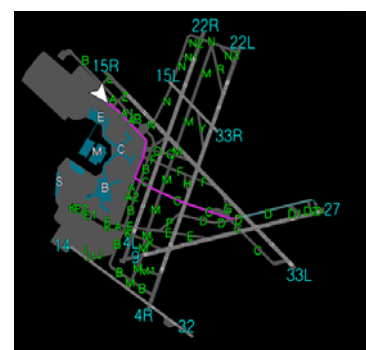


Figure 4. Taxi Message, Proposed and Accepted Route Displays

being taxing the aircraft. The flight crew in Data Comm scenarios received the route via Data Comm, but then required a voice instruction from the ground controller to begin taxing the aircraft. The voice message exchange might be: “Ground, NASA557 ready to taxi”, “NASA557, Ground, taxi via datalink route”, “NASA557, Roger.”

G. Experiment Protocol

The 22 subject pilots were scheduled and paired by flight organization to minimize potential adverse effects from differing standard operating crew procedures inherent in different airlines. During the experiment, the pilot qualified as a Captain performed the role of the PF in the left crew station and was responsible for control of the simulated aircraft throughout the experiment, and the pilot qualified as a First Officer performed the role of PM in the right crew station and had primary responsibility for Data Comm messages. All subjects were current and qualified B757/B767 pilots from five different US Part 121 Air Carriers. The 11 Captains averaged 52.5 years old with 23.9 years of commercial flying and 17,600 hours of flight time. The 11 First Officers averaged 44.2 years old with 15 years of commercial flying and 11,200 hours of flight time.

The first day’s schedule began with two hours of academic training that included an overview of the FAA NextGen implementation roadmap and the role of Data Comm, practice sending and receiving CPDLC messages, and a detailed discussion of the eight scenarios. During this training it was stressed that the crew were to operate safely and in accordance with their airline’s standard operating procedure. It was also briefed that CPDLC Data Comm messages were to be responded to by the PM (the First Officer), however the PF (Captain) could respond if that seemed appropriate for the crew. All messages were to be responded commensurate with other cockpit tasks and priorities. Training was followed by oculometer calibration and then two hours of training runs in the IFD. The afternoon consisted of eight scenarios, with the remainder of the scenarios completed the morning of the second day. Subject pilots completed an electronic post-scenario questionnaire in the IFD after each run, and completed a hand written post-experiment questionnaire in the briefing room. Both asked about crew workload, SA, and acceptability of using Data Comm in a busy terminal area, with the post-scenario focused on that particular run and the post-experiment on comparisons between scenarios.

IV. Experiment Results

A complete presentation and discussion of the results cannot be given herein, but are contained in a NASA Technical Paper (in press). A subset of the results are presented here.

A. Mean Response Time To Data Comm Messages

The mean response time from uplink of Data Comm message until crew response is shown in Table 2. This table also shows the number of occurrences (N) and computed standard deviation. The data suggests that the majority of response times were well under a minute (mean = 20.7 sec, standard deviation = 17.6 sec across all conditions).

Table 2. Mean Response Time And Standard Deviation By Condition (seconds)

Condition →	Arrival / Paper	Depart / Paper	Arrival / MMD	Depart / MMD	Arrival / MMD+Rte	Depart / MMD+Rte
Mean	18.7	22.9	17.8	21.0	18.2	24.0
Std Deviation	14.6	20.4	15.7	16.6	12.8	21.2
N	146	188	153	197	147	185

In Table 3, the mean response time differences are shown by pairwise comparison across flight phase (arrival or departure) and display condition (paper, MMD-Only, or MMD+Route). Negative differences (shown in parentheses) indicate when the column condition exhibited a mean response time which was *faster* than the row comparison. The analysis indicated that the only significant effects ($F(5,1010) = 3.777$, $p=0.002$) were between the two lowest (Arrival/MMD and Arrival/MMD+Rte) and the highest (Depart/MMD+Rte) mean response times. Arrival response times tended to be less than departure response times due to the more complex taxi routes and taxi CPDLC messages given on departure. The remaining response times by condition did not differ significantly.

Table 3. Differences in Mean Response Time by Pairwise Comparisons (seconds)

	Depart / Paper	Arrival / MMD	Depart / MMD	Arrival / MMD+Rte	Depart / MMD+Rte
Arrival / Paper	(4.2)	0.9	(2.3)	0.5	(5.3)
Depart / Paper		5.1	1.9	4.7	(1.1)
Arrival / MMD			(3.2)	(0.4)	(6.2)*
Depart / MMD				2.8	(3.0)
Arrival / MMD+Rte					(5.8)*

* Statistically significant to $\alpha = 0.05$ level

In Table 4, the mean response time and standard deviation from uplink of Data Comm message until crew response mean is shown by message type. In Table 5, the mean response time differences are shown by pairwise comparison across the message types. Negative differences (shown in parentheses) indicate when the column condition exhibited a mean response time which was *faster* than the row comparison. The analysis indicated a significant effect ($F(5,1010) = 7.602$, $p < 0.0001$), for all of the comparisons involving Expected Taxi messages. The remaining response times by condition did not differ significantly.

Table 4. Mean Response Time And Standard Deviation By Message Type (seconds)

Message Type	Info	Frequency	PB/Start	Expected Taxi	Taxi	Amended Taxi
Mean	19.0	16.0	19.5	26.5	19.8	18.6
Std Deviation	19.3	12.6	17.0	20.7	16.6	12.2
N	250	65	127	253	132	190

Table 5. Mean Response Time Pairwise Comparisons By Message Type (seconds)

	Frequency	Push back and Start	Expected Taxi	Taxi	Amended Taxi
Info	3.1	(0.5)	(7.4)*	(0.8)	0.4
Frequency		(3.6)	(10.5)*	(3.9)	(2.6)
Pushback and Start			(6.9)*	(0.3)	1.0
Exp Taxi				6.7*	7.9*
Taxi					1.2

* Statistically significant to $\alpha = 0.05$ level

96% of all CPDLC messages were replied to within 2 minutes. 34 of the 1056 Data Comm uplink messages were not responded to within 120 seconds, and 5 were not responded to at all. In the 34 cases where response time was longer than 120 seconds, it appeared to the researchers that the root cause was the pilots read the uplink message and mentally processed it, were complying if appropriate, but either believed they had acknowledged the message or forget to acknowledge the message on the second page of the FANS 1/A implementation. This researcher observation is corroborated by crew debrief comments, where they commented having to proceed to a separate page to respond led to occasional mistakes. In all 5 cases where a Data Comm uplink was not responded to, the root cause was the arrival scenario was terminated before final “Amended Taxi-In” message was acknowledged. An aggregate plot of CPDLC Data Comm message response time distributions is presented in Figure 5.

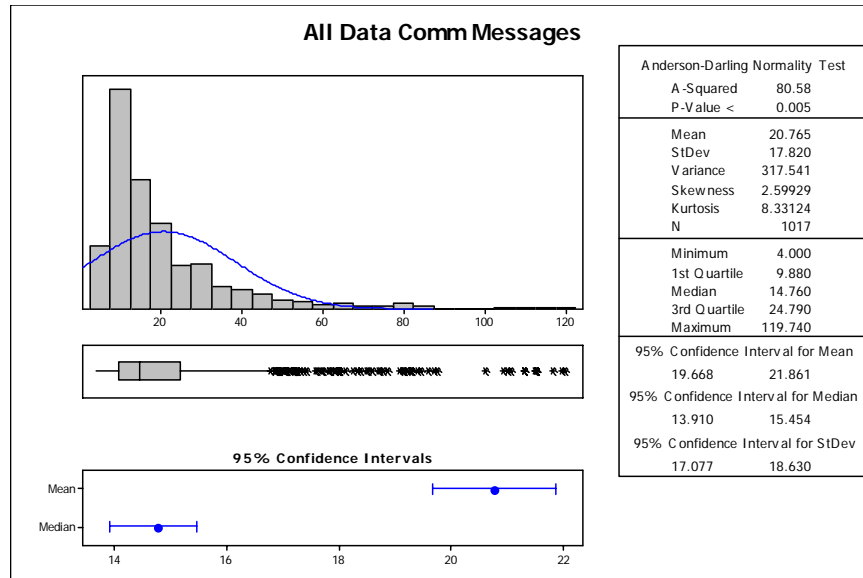


Figure 5. Response time distribution for all CPDLC messages (seconds)

B. Head-Down vs. Head-Up Time

The oculometer eye tracking data were analyzed to determine when the pilots were “head-down” or “head-up.” The data was analyzed by phase of flight and are presented in Figure 6, showing the percentage of time the PF and PM were head-up by altitude band, arrival taxi, and departure taxi. The data clearly indicated that, regardless of experimental condition, pilots focused a majority of their attention inside the cockpit while airborne. These results stand in sharp contrast to the FAA recommendation stating $\frac{2}{3}$ to $\frac{3}{4}$ of the pilot’s scan should be outside the aircraft.¹⁸

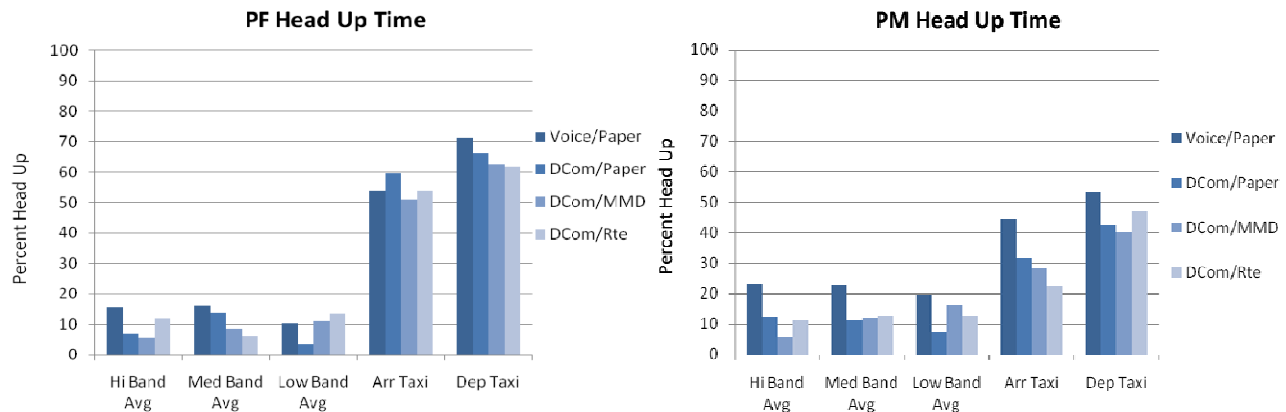


Figure 6. Percent head-up time by altitude and scenario condition

Aggregate head tracking analysis indicated that the overall effect observed was a statistically insignificant decrease in the PF head-up time in scenarios involving Data Comm, with significant difference across conditions for the PM ($F = 4.03$, $p = 0.008$). It is postulated that the increased requirement for the PM to interface with Data Comm messages in these scenarios using a CDU mounted in a relatively low location in the cockpit reduced the capacity for frequent lookout tasks. Statistical analysis comparing the variance between arrival and departure taxi scenarios indicated the departure taxi scenario yielded significantly greater head up than the arrival taxi scenario, $F(1,3) = 42.09$, $p < 0.001$. This difference is postulated due to the longer and more complex departure taxi routes. Despite the statistical (or not) differences in computed head-up time, the crew questionnaire responses indicated that the increase in head-down time associated with Data Comm employment was acceptable with the exception of from the Final Approach Fix to landing rollout during arrival, and crossing an active runway during departures.

C. Flight Crew Workload

Pilot workload was measured using a post-scenario, Bedford Workload assessment. The ordinal scale ranges from 1 (workload insignificant) to 10 (task abandoned), with rating of 3 (enough spare capacity for all desirable additional tasks) or less to be considered operationally acceptable. PF ratings of workload in flight operations were significantly higher than PM ($X^2(1)=7.794$, $p=0.005$), but were not for surface/taxi operations ($X^2(1)=2.339$, $p=0.126$). A binomial test with a cut-point of 3 (“Enough spare capacity for all desirable additional tasks”) and test proportion of 75% showed that most ratings were significantly on the low workload side of the scale for PF and PM, for both inflight and surface operations.

The left graph in Figure 7 shows the median responses for PF (in blue) and PM (in green) workload ratings for flight portions of the arrival scenarios. PF rated flight operation workload significantly differently among the display conditions ($X^2(3)=8.145$, $p=0.038$). Display conditions did not differentially affect PM ratings ($X^2(3)=5.749$, $p=0.125$). Dunnett’s C post-hoc comparisons found no significant ($p>0.05$) differences among display conditions for either subject group.

The right graph in Figure 7 shows the median responses for PF (in blue) and PM (in green) workload ratings for surface and taxi operations occurring in both arrival and departure scenarios. Both PF and PM rated surface taxi operations workload significantly different among the display conditions (PF: $X^2(3)=16.522$, $p=0.001$; PM: $X^2(3)=12.515$, $p=0.006$). Analyzing surface and taxi operations for arrival scenarios and departure scenarios separately shows that the differences found in workload during these operations were significant for both PF and PM during arrival scenarios (PF: $X^2(3)=10.342$, $p=0.016$; PM: $X^2(3)=15.459$, $p=0.001$) but not departure scenarios (PF: $X^2(3)=2.982$, $p=0.394$; PM: $X^2(3)=1.875$, $p=0.599$). Dunnett’s C post-hoc comparisons indicate that for both PF and PM, the Data Comm/Paper condition had significantly higher ratings than the Voice/Paper condition. PF also showed that the Data Comm/Paper condition ratings significantly exceeded the Data Comm/MMD+Route condition ratings.

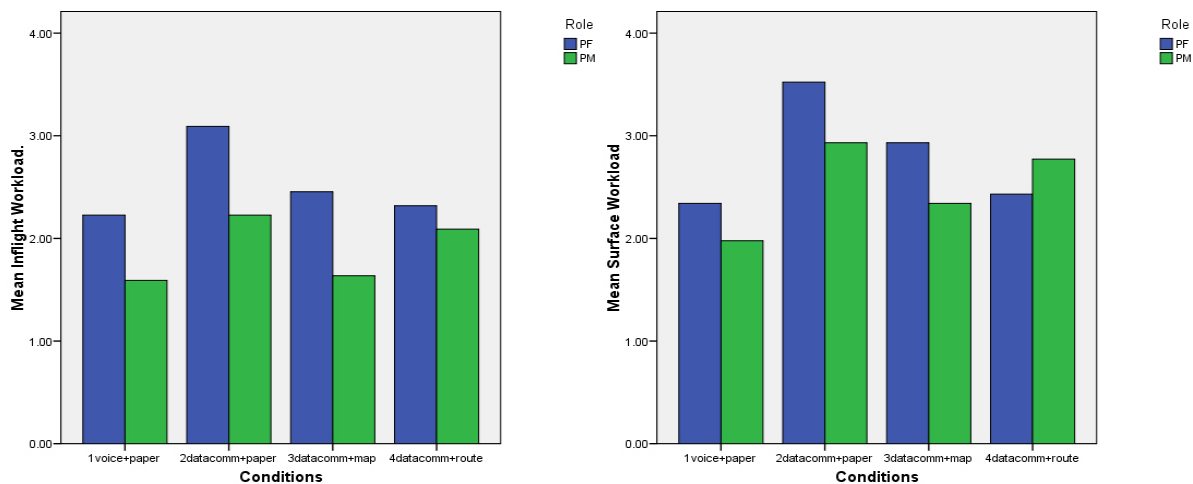


Figure 7. Mean PF and PM workload scores during inflight and surface operations

Regardless of the scenario segment, display condition, or role of the respondent, mean workload rating was never extreme. The Voice/Paper condition may have an advantage over the Data Comm/Paper condition due to familiarity, or the interface for receiving messages using Data Comm may impose additional workload. For the PF, only by adding both the moving map and route information does workload approach that of the Voice/Paper condition. The finding that PM rated workload higher for Data Comm/MMD+Route than Data Comm/MMD appears inconsistent when compared to the PF. Future research should aim to determine the contribution of these influences on perceived workload.

H. Flight Crew Situation Awareness

Situation awareness was subjectively measured using the Situation Awareness Rating Technique (SART). A binomial test of SART scales, which can range from -5 (low SA) to 13 (high SA), was conducted for both PF and PM data with a cut-point of 4 (the mid-point of the scale) and a threshold of 90%. Results indicate that the preponderance of the data for both pilot roles and for each SART measure favored the high end of the scale, demonstrating high SA on the whole. PF and PM SART ratings significantly differed for flight scenario segments ($X^2(1)=16.341$, $p<0.001$), surface/taxi operations in arrival scenarios ($X^2(1)=4.450$, $p=0.035$), but not surface/taxi operations in departure scenarios ($X^2(1)=0.872$, $p=0.351$). Results from the post-scenario questionnaire indicate SA was considered acceptable by the crews regardless of communication modality, display methodology, or the altitude at which Data Comm messages were given to the crew. Experimental conditions did not differentially affect SA ratings for either crew role during the flight scenario segments; however during surface/taxi operations in arrival scenarios, PM experienced significantly higher SA when using the Voice/Paper condition than any other Data Comm condition. PF showed this bias to Voice/Paper only over the Data Comm/Paper condition.

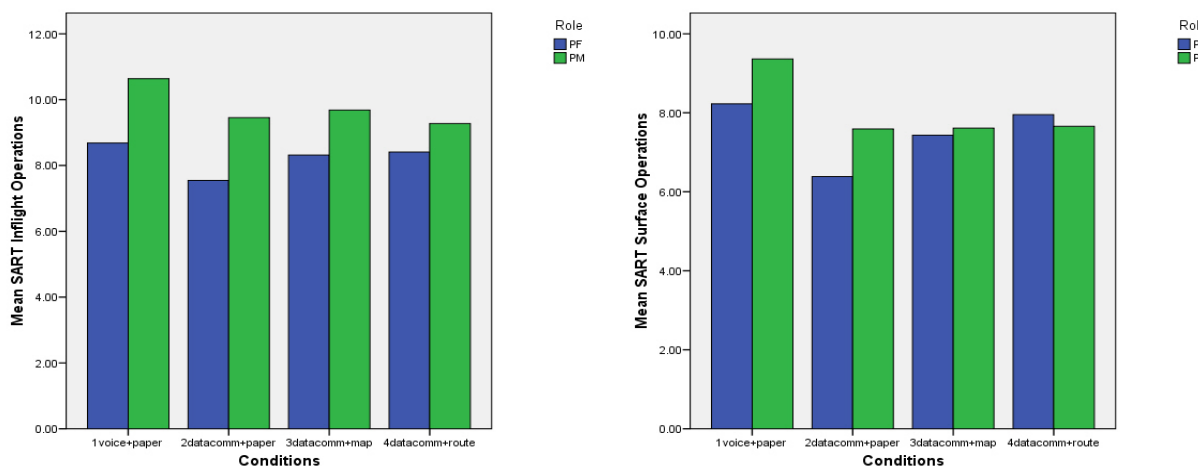


Figure 8. Mean SART scores for inflight and surface operations

I. Acceptability to Flight Crew

Post-scenario questionnaire ratings were used to assess the acceptability of Data Comm in all conditions. One overarching acceptability question (Figure 9) and eight specific questions were asked, using a scale of 1 (acceptable) to 7 (unacceptable), with a rating of 4 or less considered operationally acceptable. The PF and PM rating data are not differentiated in Figure 9, but the PFs rated the overall acceptability slightly lower than the PMs. In particular, PF ratings indicated greater concern that the taxi instructions may not be accurate, there was more operational risk, less time to respond, and less acceptable head-down time for reading and responding to CPDLC messages.

When asked to assess overall acceptability, PM ratings did not show a preference among Data Comm conditions, but PF indicated that the Data Comm/Paper condition was less acceptable than the Data Comm/MMD+Route condition. However subjects rated the Voice/Paper as more acceptable than Data Comm conditions.

Post-experiment rating responses to the acceptability of receiving Data Comm messages from controllers did not vary significantly by crew position or by display methodology. Using a rating scale of YES or NO, on average, the 22 subject pilots responded that it was acceptable for controllers to send Data Comm messages to the flight crew in a busy terminal area with two exceptions and one caveat. The exceptions of when not to send CPDLC messages were

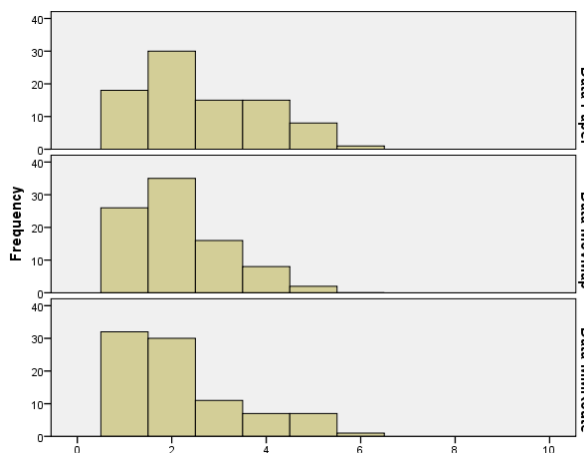


Figure 9. Overall Acceptability of Data Comm

the time between the Final Approach Fix and below 80 KIAS on landing roll-out, and while crossing an active runway during taxi operations. The caveat was if the CPDLC message was sent without a chime and the crew did not have to respond immediately, it would be acceptable to send a CPDLC message even in those instances.

From the post-experiment questionnaire, the number and percentage of crews who stated it is acceptable for a controller to send CPDLC uplink messages were:

• above 10,000' MSL:	22	100%
• between 10,000' and Final Approach Fix:	18	82%
• between FAF and below 80 KIAS on roll-out:	3	14%
• during taxi or surface operations:	21	95%

J. General Debrief Comments

Following the post experiment questionnaire, a 90 minute verbal debrief session was held with the two subject pilots and the research team. Given the assumptions of the experiment and what the crews experienced for Data Comm in a Segment 2 environment, the following miscellaneous comments regarding Data Comm were made by the crews:

- Expect the flight crew to require at least one minute to respond to ATC Data Comm uplink messages in the Terminal Area, and two minutes would be preferable.
- Many of the pilots had issues with two messages in different mode to cross active runway (timing, priority, etc). They agreed that an ATC voice instruction was essential, but most did not think it was appropriate to go head-down to acknowledge the Data Comm message. An option was to send the Data Comm message without a chime.
- Two different categories of Data Comm messages: important and informational (may require different crew response times, with / without an audible chime, etc). Important messages imply a change to the aircraft's route and a response by the crew (taxi clearance with a response within two minutes), whereas informational messages may not meet both criteria (new altimeter setting, new "Expected Taxi" clearance).
- Most crews thought "Expected Taxi" messages were useful.
- Most crews thought "Pushback" and "Start" messages should be combined.
- Many crews recommended that the ATC voice instruction to begin taxiing be given by the controller without the need for a voice request from the flight crew. It did not seem to be necessary for the flight crew to request a taxi clearance via Data Comm, then again via voice.
- All crews except one stated the downlink response should be on same page as uplink message to reduce head-down time and potential for error.
- A need was identified for a method to handle outdated messages and displays.
- "Expected Taxi" message should include hold short instructions in the text message, and red hold short bars when displayed graphically on the ND.

V. Conclusion

A. Hypothesis 1: Pilot workload and SA will differ significantly between communications modalities.

Post-scenario workload ratings remained favorable (low or adequate workload) in both voice and Data Comm scenarios. PF workload was statistically higher for Data Comm than voice, though still in the adequate region. Post experiment workload comparison preference ratings improved in the presence of Data Comm.

SART ratings indicated no statistically significant SA differences for either crew role during the flight scenario segments; however, during surface/taxi operations in arrival scenarios, the PM experienced significantly higher SA when using the Voice/Paper condition than any other Data Comm condition. The PF also showed this trend of Voice/Paper over the Data Comm/Paper condition. However, the post-scenario questionnaires indicate that SA was considered acceptable by the crews regardless of communication modality.

The amount of time a pilot devotes to being head-up can be considered as an SA indicator. Eye gaze data for both the PF and PM showed significantly more head-down time in Data Comm scenarios than in voice scenarios. This increase in head-down time was not reflected in workload or acceptability ratings but suggests that the pilot's ability to attend to head-up responsibilities might be impaired.

B. Hypothesis 2: Pilot workload and SA will differ significantly between display modes when using Data Comm.

No statistically significant difference existed in mean response times to Data Comm messages by display methodology. Post scenario workload ratings remained favorable in both voice and Data Comm modalities. PF workload was statistically higher for Data Comm than voice modality, though still in the adequate region. Post experiment workload preference ratings improved when implementing loadable routes on the MMD.

SART ratings indicated that the display modes (Moving Map and Moving Map and Route) provide improved SA over paper-only in a Data Comm environment, although the differences were not statistically significant. In all cases, the post-scenario questionnaire indicate that SA was considered acceptable by the crews regardless of the display methodology.

Head-up time, as a possible SA indicator, was not found to be significantly different between the Data Comm display modes.

C. Hypothesis 3: Pilots will rate the Data Comm used within this experiment as operationally acceptable.

Overall, low Data Comm response times were observed (a mean of 20.7 seconds). Approximately 4% of the Data Comm messages were not responded to within two minutes. Researcher observation and flight crew comments during the post experiment debrief session indicated these late responses were due to the crew believing they had responded to the Data Comm message, or forgetting to acknowledge. In all cases the message was read by the PM and briefed to the other crew member.

Statistically and operationally significantly more head-down time for both crew roles existed in Data Comm/Paper compared to Voice/Paper, with head-down time varying by display condition. This increase in head-down time was not deemed unacceptable by the crews, nor was it reflected in workload or acceptability ratings.

Post-scenario workload ratings remained favorable in Data Comm modalities for all display conditions.

All crews indicated receiving Data Comm uplink messages during approaches above 10,000' would be acceptable, with acceptability decreasing as the aircraft approached landing. Most crews indicated that Data Comm should not be used after the FAF until below 80 KIAS during landing roll-out, however many crews also stated Data Comm messages during that time would be acceptable if they were not accompanied by a chime and the flight crew was not expected to immediately respond to the message.

Crews indicated the use of voice communication to cross an active runway while taxiing was necessary and appropriate. It was also stated the simultaneous use of a Data Comm message was not appropriate since it caused crew members to go head-down at a critical time. Use of a Data Comm message without a chime was considered an acceptable alternative.

D. Summary

The experiment utilized high workload arrival and departure scenarios to study the impact to flight crews using Data Comm in a busy terminal area. The crews' qualitative comments indicated general, strong acceptance of employment of Data Comm in the terminal area within the scope of this experiment, and in terms of workload, a strong preference for Data Comm employment. The crews rated the Expected Taxi message as useful in both arrival and departure scenarios. Qualitative data showed that crews did not find the increase in head-down time associated with Data Comm scenarios unacceptable. There was a strong desire to limit Data Comm during certain critical phases when it was important for the crew to be head-up, in particular after the Final Approach Fix until slowing below 80 knots during landing roll out, and any time crossing an active runway. However, the crews also stated that changes in the Data Comm interface to improve the ease of responding and reducing the intrusiveness of chimes and alerts were potentially mitigating the unacceptability of Data Comm messages in even these segments.

Acknowledgments

The authors greatly appreciate the vital software programming of Paul Sugden, and the work of Dennis Frasca, Jerry Karwac, Wendy Pifer, Brittany Williams, Brian Hutchinson, Darrell Sacra, Phil Smith, and Tom Feigh.

References

¹ *Concept Of Operations for the Next Generation Air Transportation System*, Version 3.0, Joint Planning and Development Office (JPDO), Washington DC, 1 October 2009

- ² Navarro, C., Sikorski, S., *Datalink Communication in Flight Deck Operations: A Synthesis of Recent Studies*, University of Toulouse, The International Journal of Aviation Psychology, pg 361-376, 1999
- ³ Rehmann, A., *Flight Simulator Evaluation of Baseline Crew Performance with Three Data Link Interfaces*, FAA Technical Center, DOT/FAA/CT-TN95/19, Sept 1995
- ⁴ Pinska, E., Whiteley, M., *LINK 2000+, France Real-Time Simulation Project*, Eurocontrol Experimental Center, EEC Report No. 395, Sept 2004
- ⁵ *D-TAXI Trial Final Report*, Version 2.0, April 2007, CASCADE, NLR, et al
- ⁶ Roeder, M., *EMMA Publishable Final Activity Report*, DLR, D013, 2007
- ⁷ Jakobi, J., *EMMA Recommendations Report*, DLR, Document D6.8.1, 2007
- ⁸ Jakobi, J., *EMMA2 Recommendations Report* (European Airport Movement Management by A-SMGS, Part 2), DLR, Document No. 2-D6.7.2, Version 1.0, 2009
- ⁹ Jakobi, J., Teotino, D., Montebello, P., *Higher-Level Services of an Advanced Surface Movement Guidance and Control System (A-SMGCS)*, ATC Quarterly, in press
- ¹⁰ Waller, M., Lohr, G., *A Piloted Simulation Study of Data Link ATC Message Exchange*, NASA TP-2859, Feb 1989
- ¹¹ Scanlon, C., Know, C., *Flight Test Show Potential Benefits of Datalink as Primary Communication Medium*, SAE Aerospace and Technology Conference, 1990
- ¹² Knox, C., Scanlon, C., *Flight Tests With a Data Link Used for Air Traffic Control Information Exchange*, NASA TP-3135, Sept 1991
- ¹³ Hooley, B., Foyle, D., Andre, A., Parke, B., *Integrating Datalink and Cockpit Display Technologies into Current and Future Taxi Operations*, Proceedings of the AIAA/IEEE 19th Digital Avionics System Conference, 2002
- ¹⁴ Parke, B., Renfro, D., Kanki, B., Hooley, B., Munro, P., Patankar, K., Foyle, D., *The Effects of Advanced Navigation Aids and Different ATC Environments on Task-Management and Communication in Low Visibility Landing and Taxi*, 11th International Symposium on Aviation Psychology, 2001
- ¹⁵ Mueller, E., *Experimental Evaluation of an Integrated Datalink and Automation-Based Strategic Trajectory Concept*. 7th AIAA Aviation Technology, Integration and Operations Conference (ATIO), Paper No. AIAA 2007-7777, Belfast, Northern Ireland, 18-20 Sept 2007
- ¹⁶ Mueller, E., and Lozito, S., *Flight Deck Procedural Guidelines for Datalink Trajectory Negotiation*. The 26th Congress of International Council of the Aeronautical Sciences (ICAS), Paper No. AIAA 2008-8901, Anchorage, Alaska. 14-19 Sept 2008
- ¹⁷ Roscoe, A.H. and Ellis, G.A., *Subjective Rating Scale for Assessing Pilot Workload in Flight: A Decade of Practical Use*. Royal Aeronautical Establishment, Report No. BR-114794, Farnborough, United Kingdom, Mar 1990
- ¹⁸ FAR/AIM, Federal Aviation Regulation/Aeronautical Information Manual, Section 8-1-6(c), *Scanning for Other Aircraft*, 2010
- ¹⁹ Taylor, R.M., *Situational awareness rating technique (SART): The development of a tool for aircrew systems design*. Proceedings of the AGARD AMP Symposium on Situational Awareness in Aerospace Operations, CP478. Seuil-sur Seine: NATO AGARD. 1989.