Evaluation of Mixed-Mode Data-Link Communications for NextGen 4DT and Equivalent Visual Surface Operations

Lawrence (Lance) J. Prinzel III, Kevin J. Shelton, Denise R. Jones, Angela S. Allamandola, Jarvis (Trey) J. Arthur III, and Randall E. Bailey

NASA Langley Research Center, Hampton, VA, USA 23681

By 2025, U.S. air traffic is predicted to increase 3-fold and may strain the current air traffic management system, which may not be able to accommodate this growth. In response to this challenge, a revolutionary new concept has been proposed for U.S. aviation operations, termed the Next Generation Air Transportation System or "NextGen". Many key capabilities are being identified to enable NextGen, including the use of data-link communications. Because NextGen represents a radically different approach to air traffic management and requires a dramatic shift in the tasks, roles, and responsibilities for the flight deck, there are numerous research issues and challenges that must be overcome to ensure a safe, sustainable air transportation system. Flight deck display and crew-vehicle interaction concepts are being developed that proactively investigate and overcome potential technology and safety barriers that might otherwise constrain the full realization of NextGen. The paper describes simulation research examining data-link communications during 4DT and equivalent visual surface operations.

INTRODUCTION

Next Generation Air Transportation System

NextGen represents a revolutionary approach to air traffic management that requires a shift in the tasks, roles, and responsibilities for all operators, including the flight deck. NextGen will remove many of the constraints in the current air transportation system, support a wider range of operations, and deliver an overall system capacity up to 3 times that of current operating levels. New capabilities are envisioned for NextGen, including four-dimensional trajectory (4DT) operations, equivalent visual operations (EVO), super density arrival/departure operations, and network-centric operations. As part of this work, NASA is evaluating the issues associated with emerging NextGen operational concepts in the Terminal Maneuvering Area (TMA) and the effect of changing communication modalities to create a net-centric environment and the impacts from increased pilot responsibility for self-separation and navigation performance compliance.

Voice Communications

Verbal communication is critical in today's National Airspace System. It's the most efficient and adaptable form of communication and has inherent advantages including speed of transmission, human tone to express urgency and intention, flexibility of dialogue, and provision of party-line or broadcast effect. Pilot-Air Traffic Control (ATC) aural communications have historically been quite successful yet are far from error-free. One content analysis showed that 40% of ATC-pilot communications contained at least one irregularity [1], but the error rate was less than one percent [2]. Fortunately, 60 to 80% of these errors were caught in the read-back process. However, the proportion of corrected read-back errors varied, where the highest workload sectors had the least corrections: en route controllers corrected 89% of the read-back errors, TRACON and local controllers caught only 60%, and 63%, respectively, and only 50% of the read-back errors on the ground frequency were corrected.

Problems with language interpretation, clarity, radio congestion (blocked transmissions), confusion, read-back errors, message factors (formulation, packaging, and delivery), and extraneous environmental noise can lead to misunderstandings during voice message reception [3]. Cushing [4] categorized these into groups involving communication problems based on language and those based on non-language. Language-based communication problems include ambiguity, homophony, punctuation, intonation of speech, speech act, problems with reference (including uncertain reference, uncertain addressee, and unclear hand-offs), problems with inference (including implicit and lexical inferences, unfamiliar terminology, and false assumptions), and problems involving repetition (such as no read-back, partial read-backs, ritualization of speech, and cultural issues). Non-language communication problems include issues with numbers (including range overlaps and transposition) problems with radios (including loss of radio and misuse of radios), and problems with compliance (such as distractions, interruptions, fatigue, non-cooperation, etc.).

Data-Link Communications

The Need. The rapidly decreasing availability of radio-frequency bandwidth is driving a rapid shift from voice to data-link. The EuroControl/FAA Future Communications Study Operational Concepts and Requirements Team stated, "The growth in aviation traffic has resulted in enormous pressure on the ability of existing spectrum resources to satisfy communications requirements. The very high frequency (VHF) spectrum is already congested and will be the significant limiting factor in Air Traffic Management (ATM) capacity by 2015 unless additional air/ground communication sources are made available to supplement the existing system" [5]. By the year 2015, only 60% of communications are projected to be provided via voice during the most critical phases of flight in the Airport/Terminal Maneuvering Area (TMA). By 2030, 85% of Air Traffic Services communications are projected to be provided via data-link in the TMA environments [5].

The Benefits. A key enabler of operations envisioned under NextGen is data-link communications and a voice-by-exception (i.e., data-link for routine communications) airspace environment. Data-link communications are a critical determinant to the expansion of air traffic capacity and to the success of an adaptable, "net-centric" operational capability under NextGen. As presently envisioned, the data-link interface changes the modality of pilot-ATC communications from aural to visual using text read-outs on the Flight Management System (FMS) Control-Display Unit. With this modality shift, research shows data-link communications generally provide the following positive benefits for pilot-ATC communications [6 – 13].

- Reduction or elimination of message blocking and congestion
- Higher efficiency and capacity of communications system resulting in improvement of message delivery time (i.e., when compared to current limited radio-frequency bandwidth, stepped-on messages, and controller-pilot read-back times)
- Persistence of the message
- Unloaded memory burden from lengthy messages, and ability to review later
- Improved information-processing efficiency and accuracy
- Possibility of effective multitasking due to user-pacing communication tasks and elimination of continuous listening workload
- Ability to enable automatic gating that may improve information transfer to other ATC and flight deck subsystems
- Inherently easier and more accurate communications for crews with language and cultural barriers

The Issues. While the transition to data-link communications would appear to be extremely beneficial, there are also many issues/concerns to be addressed. Research has shown that:

- Keyboard FMS implementations might be advantageous to non-English-native languages and cultures, but native users find the interaction to be cumbersome and reading is slower compared to aural/oral communication
- Workload and increased "head-down time" are a major concern [6] [14]
- Users are not able to gain a sense of urgency when using text instead of voice. Data-link does not adequately convey urgency and other natural voice information [15] [8]
- Users may be deprived of "party-line" information [18]
- Crew coordination, diffusion of responsibility, and reduced cross-check involving automation can occur, particularly when data-link is "message-gated" and passed directly into FMS without the pilot reading the message and entering the data manually [15] [9] [16] [17]
- Data-link changes the information processing requirements in the cockpit, requiring more manual interaction and fewer verbal responses than radio, and is more susceptible to interference and decay

Research Need. Navarro and Skiorski [19] detailed the advantages and disadvantages of data-link for today's commercial operations. Although 10 years old, the review documented the prevailing issues that continue to surround the use of data-link implementation. The authors suggested several areas that need more study to include:

- Risk of crew exclusion from communication loop between ATC and aircraft
- Workload modification due to manual and automatic data transfer between the data-link interface and FMS
- Transaction time variation due to manual and automatic data transfer between the data-link interface and FMS
- Risk of interference between simultaneous communications and flying actions
- Possibility of reducing errors by facilitating retention of information (e.g., messages kept on display)
- Possibility of reducing errors by using appropriate presentation modes of messages (graphic, iconic, or textual) and presentation formats of these messages (readability)
- Facilitating reaction time by adequate coding for urgent messages
- Increasing intra-cockpit communication (to ensure shared representation of flying conditions)
- Increasing visual tasks and visual channel overload
- Risk of reducing or removing party-line information

These research areas remain today as significant challenges to data-link and mixed-voice communications operations. Data-link may improve one source of miscommunications – the inability to get the message from one party to the other – but it does not necessarily address the rest of the communications process – i.e., whether the message was understood and whether it accurately conveyed the sender's intent. These data-link issues may be more acute for NextGen - with closer spacing and more pilot responsibility for separation – especially where increased head-down time and increased workload could significantly reduce safety margins.

Flight Deck Display Solutions

The challenge for NextGen is to replicate the strengths and advantages of the previous communication method (i.e., voice) while bringing to bear only the benefits of the new. NASA is conducting a research portfolio which strives to meet this goal. Part of this work targets flight crew assessment of the modality change effects associated with data-link operations, the impact of party-line information changes, and potential mitigating or confounding factors for NextGen operations. Past research has demonstrated that the

efficacy of party-line information is dependent upon crew resources, communication, and workload. These influences are undoubtedly also impacted by auditory demands and other visual data available to the flight crew. In particular, what potential exists in the information available by use of ADS-B "in" capability, with traffic intent data? Could the visual display of data-link information and associated navigation and intent data serve to replace voice (voice-by-exception) for routine communications and compensate for the potential impacts of party-line information loss or inherent latencies in managing mixed voice and data pilot-ATC communications with Controller-Pilot Data Link Capability (CPDLC).

To achieve the NextGen vision, research is necessary to proactively investigate technology which may overcome potential safety barriers that would otherwise constrain surface traffic optimization, maximum runway capacity through reduced runway occupancy time, simultaneous single runway operations, 4DT design and operation, and equivalent visual operations, among others [35, 41]. These safety objectives were the focus of a research study conducted at NASA Langley Research Center that attempted to examine a set of safety issues associated with data-link communications with surface operations being envisioned for the Next Generation Air Transportation System.

NASA NextGen Data-Link Communications Research

The present study evaluated how data-link implementations (i.e., redundant, supplementary, or replacement of voice) and mixed-modality communication (i.e., voice and data-link) impact crew decision-making, workload, situation awareness, and crew coordination during operational situations conducted under envisioned 4DT EVO NextGen operational concepts. These four communication modes represent possible communications concepts during TMA operations being considered for NextGen or its developmental or evolutionary instantiations: (a) voice with redundant data-link (data-link as redundant to voice), (b) voice and data-link (data-link as supplement to voice), (c) data-link only (data-link as replacement for voice), and (d) data-link only with intent. This paper specifically describes only the off-nominal performance aspects of this research. Off-nominal scenarios (e.g., incorrect clearances, pilot blunders/errors, and data-link transmission errors) were unexpectedly introduced as critical tests of the robustness and safety of the operations.

METHODOLOGY

Pilot Participants

Twenty-four pilots served as participants for the research, representing twelve flight crews. Each flight crew flew for a major U.S. air carrier and was paired by airline to ensure crew coordination and cohesion with regard to terminal and surface operational procedures. The Captains had an average of over 14,000 flight hours with 19 years of commercial flying. The first officers had an average of over 10,000 flight hours with 13 years of commercial experience. The Captains were recruited on the basis of head-up display experience and all pilots were required to have "glass cockpit" (ICAO aero-plane design code D-E) experience. The flight crews were provided with a forty-five minute briefing to explain the display concepts and evaluation tasks. After the briefing, a 1-hour training session was conducted to familiarize the flight crews with the simulator, ORD airport, published arrivals and departures, display concepts, and evaluation tasks. The Captain was the designated pilot flying throughout all the trials and the First Officer served as the monitoring pilot (pilot-not-flying). The flight crews were asked to maintain good crew resource management and comply with company specific procedures for low visibility operations.

Experimental Conditions

The research was designed to evaluate possible evolutionary implementations for NextGen concepts (circa 2008 through 2025) and their potential applications of data-link based on current (e.g., EuroControl

LINK2000+ Maastricht Upper Airspace Control Center) and future concepts and operations. To reflect evolutionary applications of data-link and CPDLC, the research examined four communication modes that have been identified to represent the gamut of potential technological implementations:

- Data-Link + Voice: The voice + data-link as <u>redundant to voice</u> condition required the pilot-not-flying to respond via both voice and data-link to ATC clearances and all ATC clearances were provided by both voice and CPDLC and heard on the VHF ("party-line") broadcast. This condition retains the benefits of voice and adds the capability for retention and review of the message in the cockpit. However, this method may significantly increase the workload of ATC and does not reduce the VHF bandwidth.

- Data-Link + Read-back Only: The voice + data-link as <u>supplement to voice</u> condition stipulated that pilots provide radio read-backs on the VHF channel of received data-link messages. ATC sent all clearances via data-link messages and all flight crews were required to respond through CPDLC and voice represented a condition in which CPDLC was supplemental to voice. This condition reduced the VHF chatter by 50% (except for ATC non-routine communications) and retains the retention/review of the message in the cockpit. This method also retains the "read-back" as a checking mechanism as to whether the communication was correctly received and understood. The disadvantage of this method is that the VHF bandwidth is reduced but possibly not sufficiently, and ATC must use both voice (listening) and data-link concurrently.

- Data-Link Only: The data-link only condition (i.e., data-link as <u>replacement for voice</u> or voice-by-exception) had no ATC clearances or pilot read-backs over the voice channel and the participants were required to respond only through data-link (except for exigent situations or non-routine communications). This technique provides the maximum VHF bandwidth reduction and simplifies the communication to a single modality. However, this condition also represents the loss of the auditory communication benefits (party-line, intonation, stress, etc.)

- Data-Link + Intent: The data-link only + traffic intent condition was identical to the data-link only condition (i.e., data-link as <u>replacement for voice</u>) but presented additional <u>traffic intent</u> information (e.g., selected traffic route and destination information). The condition was introduced to determine whether providing traffic intent to the flight crew, shown on the surface map display (see below), may mitigate safety issues described by other research studies as a consequence of reduced voice communications and loss of party line information.

Research Simulator

This research was conducted in the Research Flight Deck (RFD) simulator at NASA Langley Research Center (Figure 1). The simulated aircraft dynamics model was a medium to long-haul commercial passenger aircraft classified as an ICAO Aero-plane Design Code D. The RFD configuration is a fixed-based, dual-pilot simulator with a collimated 200° panoramic out-the-window scene. Operations were conducted at the Chicago O'Hare airport. The out-the-window scene included realistic taxiways and runways with appropriate markings, airport lighting, and other aircraft in various simulated weather/lighting conditions. The visual acuity of the out-the-window scene was 20/40 (pixels per degree). The RFD was equipped with a 30° H x 24° V Rockwell Collins HGS-4000© Head-Up Display (HUD) located on the left or captain's side.

Figure 1 shows the the simulator which had four large main instrument panel displays referred to as: (left to right) Pilot's Primary Flight Display (PFD), Pilot's Navigation Display (ND), Co-pilot's ND, and Co-pilot's PFD. The four display panels are 17in. LCD displays with 13.25 x 10.5 in. viewable space at 1280x1024 (SXGA) resolution and a 5:4 aspect ratio. These displays provided a total of 53 x 10.5 in. of display space at

5120 x 1024 pixels. The pixel densities of the main instrument panels are approximately 100 dpi. Each display has independent brightness control and has very high readability, contrast ratio and viewing angle specifications resulting in excellent cross cockpit readability. The four main displays are driven by dedicated PC-based computers with high-end graphics. The RFD includes a Mode Control Panel (MCP), Flight Management System (FMS), Control Display Units (CDU), and hydraulic side-stick control inceptors (Figure 1).

Two commercial Class III Astronautics© Electronic Flight Bags (EFBs) were installed with a display resolution of 1024 x 768 (XGA) on at 10.4 in. diagonal touch screen display. The EFBs are mounted above and outboard of the side-stick control inceptors near the side window sill. The EFBs utilize a dynamic menu interface using the line select buttons or touch-screen operation. Custom software was developed to provide desired experimental functionality, including CPDLC, airport maps and charts. The airport map page and charts pages displayed standard FAA diagrams that enabled the pilots to zoom and pan utilizing available buttons on the interface.
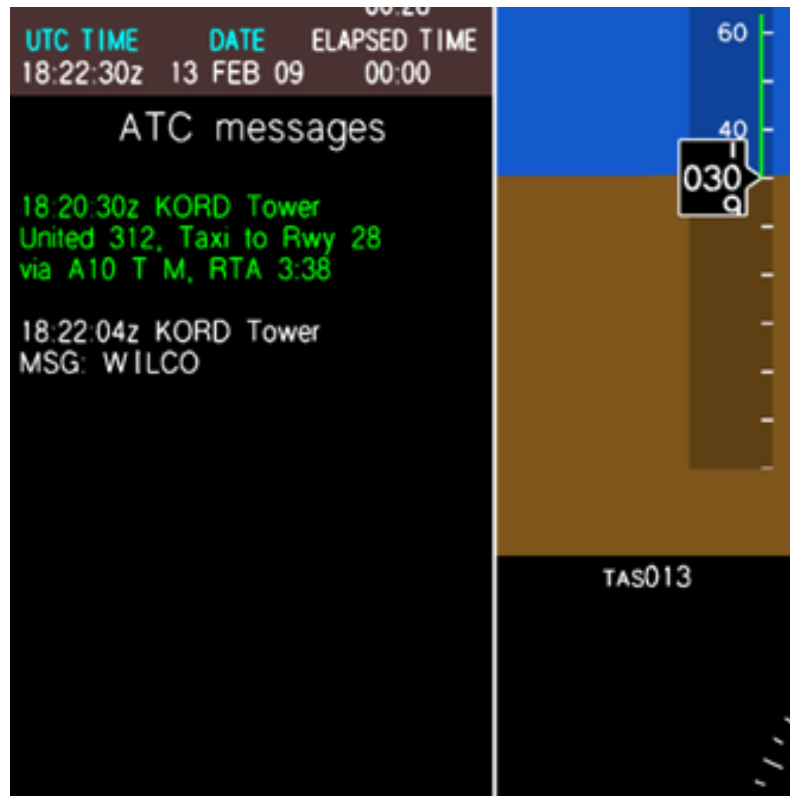
Controller Pilot Data Link (CPDLC)

Air Traffic Control was simulated and broadcasted via a standard voice audio instruction (pre-recorded audio clearances and exchanges) and a simulated data-link message based on ICAO Doc9705 compliant ATN/CPDLC technology using D0-258A message format sets. The ATC messages were triggered based on specific events and timings to coincide with the task scenario. During the trials, automated ATC communications were played to simulate the VHF communication environment for the appropriate experimental condition.

Figure 1. Research Flight Deck Simulator at NASA Langley Research Center

Flight Deck Displays

Primary Flight Display. The Primary Flight Display (PFD), Navigation Display (ND), and Engine Indication and Crew Alerting System display (EICAS) were modeled after Boeing 787, Gulfstream G550 and Airbus A350 instrument panels as current state-of-the-art production aircraft. The PFD display unit includes an ATC message area (CPDLC message area; Figure 2) on the outboard third of the display unit showing incoming and outgoing ATC data-link communications in alpha-numerical format. Incoming messages are color-coded green while outgoing messages are white. All messages are time-stamped. The captain's inboard display unit shows navigation and EICAS displays.

Figure 2. Primary Flight Display Controller Pilot Data-Link (CPDLC) Message Area

Navigation Display. The first officer's inboard display is a full-screen moving map navigation display with high airport surface detail. Unlike today's equipage, surface and airborne traffic icons were included to simulate an ADS-B-In environment. This display automatically transitions to the surface map display (described below) after landing, when the ground speed is less then 80kts (Figure 3).

Electronic Flight Bag Display. The Electronic Flight Bag (EFB) is used as a flight crew's interface for ATC data-link communications (the FMS CDU may also be used), and for approach chart and airport diagram chart references. The CPDLC / ATC functionality was modeled on existing Boeing commercial aircraft vendor interfaces.

Head-Up Display. The HUD surface operations concepts evolved from Taxiway Navigation and Situation Awareness ("T-NASA") research [21] [22] and Runway Incursion Prevention System. The T-NASA HUD

display concepts show current ground speed in digital format, the current taxiway, next cleared taxiway, and centerline markers and virtual cones on the taxiway edge for the cleared route. Additional non-conformal and conformal cues are given for pilot to assist in steering.
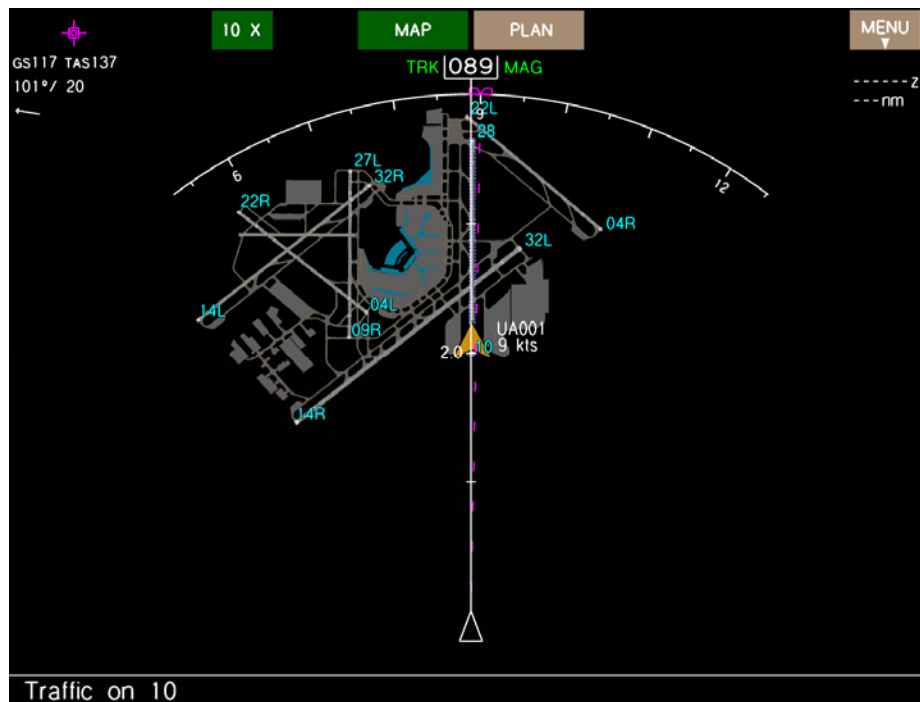
Figure 3. Navigation Display Concept Showing Airport Map

HUD format modifications are being developed to accommodate emerging NextGen operations. The surface mode of the HUD depicts 4DT guidance that utilizes dynamic algorithms to generate speed- or time-based taxi clearances to calculate the most efficient movement of all surface traffic and enable precise surface coordination [24] [25]. Current instantiations of HUD-displayed 4DT guidance symbology use speed and time commands to the flight crews throughout the taxi route to regulate the required timing precision of surface traffic movements. The aircraft's taxi speed may be adjusted if the pilot is unable to conform to the speed command to meet the required time-of-arrival (TOA), if traffic is unable to comply (creating a reduction in separation), or if the needs of the dynamic airport surface operation requires adjustment.

Surface Map Display. The surface map display is an enhanced version of the track-up navigation display. The surface map is automatically displayed on the ND during the landing rollout or it can also be manually selected at any time. The surface map display design is based upon current guidance from RTCA SC-183 working group findings in regard to element shape and color assignments, and is designed for strategic use for surface operations (to minimize head-down time). The traffic icons are medium tan colored chevrons when on the surface and cyan when airborne. Own-ship icon is a large white chevron, and the cleared taxi route is shown in magenta.

The surface map display concept contains a number of significant enhancements to help enable 4DT compliance and situation awareness. These included the alphanumerical display of 4DT time and speed guidance and distances to major taxiways and runways. An additional enhancement was a list of aircraft position and associated state data (obtained from a modeled ADS-B, TIS-B, and CPDLC-all surveillance source). This traffic list on the right side of the display is shown on the right side of the display, sorted by

proximity to own-ship with the closest traffic listed first. A cursor control device can be used to select specific traffic in the list (magenta box indicates selected traffic) which brings up additional details of the selected traffic (lower right of Figure 4). Traffic details include: type, flight ID, speed, heading, ATC assigned route (intent), and the range / bearing from own-ship position. The assigned taxi route is displayed on the lower center of the display as a text string, and current position along that route is highlighted in magenta. Distance to the next route segment is also displayed nearby in green text. 4D trajectory guidance information (textual) is displayed as: current state information in white [ground speed, elapsed time, and Estimated Time-of-Arrival (ETA) UTC]; while guidance information [Required Time of Arrival (RTA) in Greenwich Mean time (UTC), Required Time En-route (RTE) in min:sec, and guidance ground speed] is displayed in magenta.
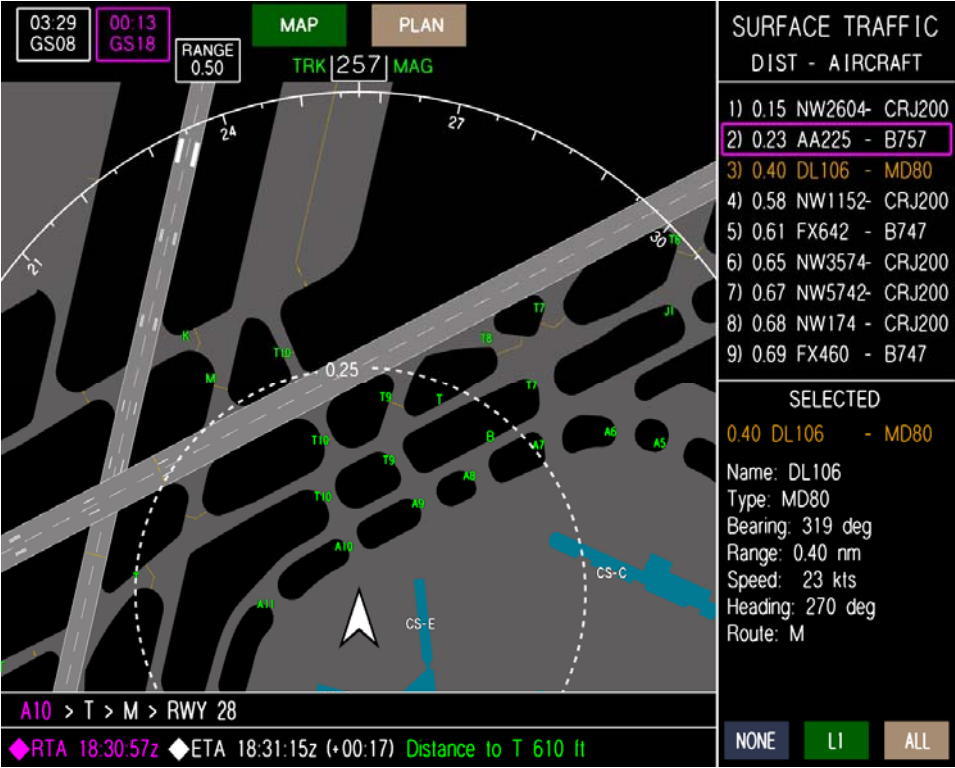

Figure 4. Surface Map Display Concept

Figure 4 shows the surface map concept (excluding traffic intent graphical presentations not used in current study) that was used for the current research to represent a current state-of-the-art baseline configuration. Shelton, Prinzel, et al. [46] contains a description of the NASA surface map display concept which includes graphical 4DT guidance, traffic, CPDLC, graphical traffic intent and state, own-ship magenta route (based on data-link ATC clearance), conformance monitoring and adherence, and other features designed to support future surface operations.

Nominal Evaluation Tasks

Pilots conducted landing, departure, and taxi operations at Chicago O'Hare International Airport (FAA identifier: ORD). A total of 24 different scenarios were used in the study that were randomly assigned and presented. The tasks were designed to emulate Next Generation Air Transportation System operations including high tempo traffic density and patterns and data-linked (CPDLC) landing and taxi clearances prior to the final approach fix (JPDO, August 2007). The weather state for the out-the-window scene represented

Category II (CAT II) landing minima weather with 1200 runway visual range (RVR) and a 50 feet ceiling. Winds were varied slightly in speed and direction to increase the simulation realism and pilot workload.

The nominal tasks were either: (a) an approach and landing with taxi to the movement area or (b) push-back and taxi to departure. Flight crews were briefed prior to each trial on the task (e.g., ILS 14R approach), weather, winds, and other relevant information required to complete the task. Because the approach trials began inside the initial approach fix, pilots were provided adequate time to review charts and brief the approach prior to beginning these data runs. These trials began upon flight crew acceptance of the CPDLC data-linked clearance which contained both the landing and taxi clearance. At that time, auto-land was engaged for the approach by selecting the mode control panel selections. Auto-throttles were also engaged to maintain 138 knot approach speed. Upon weight-on-wheels, the autopilot was manually disengaged and the Captain performed a high-speed turn-off and taxi to the airport movement area. For the departure tasks, the crew was provided with a data-link clearance for taxi to the departure runway. These trials also began upon flight crew acceptance of the CPDLC data-linked clearance. Subsequent clearances included "taxi into position and hold" with a subsequent "cleared for takeoff" clearance to fly runway heading to a specified altitude.

The subject crews were briefed to adhere to commanded 4DT guidance displayed on the surface map and HUD. These commanded speeds and RTAs faster than typical taxi speeds under low-visibility conditions (1200 RVR) as they emulate an envisioned equivalent visual operational surface concept under NextGen. Crews, however, were instructed that the safety of the aircraft should never be compromised and to taxi at speeds that they felt comfortable and safe and to notify ATC if they could not adhere to the RTA taxi clearances.

Off-Nominal Evaluation Tasks

Four off-nominal scenarios were tested. These scripted scenarios simulated the impact of human errors or blunders which caused incorrect clearances, created runway incursions, and generated data-link malfunctions. The objectives of the scenarios were to evaluate whether the flight crews would recognize off-nominal situations (which might become hazardous) and how they would respond to these events with the various data-link communication conditions (described above). The scenarios were introduced in a planned sequence using a matrix ordering on the basis of "intrusiveness" to minimize the impact of the current off-nominal event on future off-nominal detection (i.e., to avoid a confound on later trials).

Erroneous Data-link Clearance. The first off-nominal scenario was created by an incorrect data-linked clearance, that produced a route disconnect in the taxi path from Runway 14R with a high-speed turn-off onto taxiway "Mike 7". The data-link clearance included "Mike Four" after "Foxtrot" in the textual data-link message which would result in the aircraft turning back toward the active runway instead of toward the concourse creating a potential runway incursion situation. The data link clearance was in dissociation with the voice clearance (for those conditions that had voice) in that the clearance was stated correctly by the controller over the voice channel. Therefore, the flight crew could detect the anomaly if they heard and understood the dissociation between the voice and data-link clearances or by reviewing the data-link clearance fully prior to WILCO (or later during taxiing). It was hypothesized that voice (as redundant to data-link) would result in greater awareness and likelihood of detection of the error because the flight crew would have more opportunities to realize the error.

Data-Link Clearance Received for Other Aircraft. The second off-nominal scenario examined the effect of erroneously receiving a data-link clearance intended for another aircraft. The scenario began with an approach to Runway 22R. The data-link clearance received prior to the final approach fix contained the

proper landing and taxi clearance including RTA. The flight crew acknowledged the clearance through CPDLC and continued the approach including configuring the aircraft for landing and completing the landing checklist. Approximately 0.5 nm from the runway threshold, the aircraft received a second landing clearance that was erroneously sent to the aircraft although it was intended for another aircraft conducting operations on a proximate runway. The clearance contained a similar taxi clearance as the previous CPDLC message but also had errors that would result in a route dissociation between the landing runway and the taxi instructions. The erroneous message was hypothesized to be generated by a controller error where the message was sent to and contained a transposed aircraft call sign (which was similar to the own-ship aircraft). For those conditions that included voice, the voice clearance was accurate and correctly stated the intended aircraft call sign and taxi clearance. The hypothesis was that, during the high workload phase of the approach when the incorrect clearance was received, the flight crew would not be able to attend to the CPDLC message and would accept the clearance without review. This contrasts with the flight crews who also heard the voice clearance, thus, providing them an additional opportunity to recognize that the taxi clearance was not congruent with the landing runway and was not intended for their aircraft.

Taxiway Conflict. The third off-nominal scenario evaluated whether the flight crew would be task-captured by the 4DT task. The 4DT information was presented head-up and head-down on the surface moving map display, with the associated requirement to adhere to a 4DT RTA, and the scenario evaluated if the flight crew would be able to detect an aircraft that mistakenly was mistakenly taxiing in the opposite direction on the same taxiway. The scenario began as a taxi out task in which the aircraft received push back and taxi clearance to the departure runway with RTA instructions. Once the aircraft accepted the clearance, the aircraft began the taxi to the Runway 4R departure runway. Another aircraft subsequently landed on Runway 10 and turned off on taxiway Mike-7 with a taxi clearance to Delta and Alpha-17. The conflict aircraft paused at Mike-7 and was communicating, on tower frequency, that they were unsure where they were. The conflict aircraft mistakenly turned on the first left taxiway (Mike) instead of the second taxiway (Delta) resulting in a potential nose-to-nose situation. The objective of the scenario was to determine if any of the experimental conditions would result in significant head-down time and increased intra-cockpit communications that may prevent the timely identification of the other traffic (the scenario was not intended to result in a collision but would have resulted in a nose-to-nose situation requiring ground vehicles to tow the aircraft if not detected in an adequate amount of time). The hypothesis was that the flight crew that had either of the voice and data-link conditions would more easily recognize the potential conflict and would be proactive in avoiding the situation because of the reduction in hypothesized intra-cockpit communication and reduced head-down time. The flight crews that had only data-link (e.g., data-link only and data-link + intent) would instead be reactive in detecting the situation because of increased communication and head-down time involved in interacting with the EFB and surface map displays. These crews would avoid a collision but would result in both aircraft being blocked on the taxiway without an ability to turn off the taxiway. The consequence of which would be the requirement to have a vehicle tow one of the aircraft, resulting in significant reduction of operational efficiency.

Approach Runway Incursion. The final off-nominal scenario created a potential runway incursion in which an aircraft erroneously taxied onto the runway and held despite it being cleared for take-off while the subjects were on final approach to the same runway. The blundering aircraft had responded to ATC, but did not <u>correctly</u> acknowledge the clearance, and the controller failed to recognize the error (expectancy heuristic). The aircraft consequently remained on the runway, unaware that they were supposed to depart because of the aircraft on approach behind them. In this case, it was hypothesized that having party-line information would cue the flight crew that an aircraft was holding on the landing runway. Also, having data-link <u>with</u> traffic intent might also alert the flight crew to the runway hazard.

Experimental Design

The results presented herein were part of a larger research study consisting of three experiments that were conducted over two-days of testing for each flight crew. The experiments ran sequentially and the experiment described here was presented first in order. The experimental design was a mixed-subjects factorial design. The independent variables of interest were communication mode for nominal (within-subjects) and off-nominal trials (between-subjects). All flight crews experienced each of the four off-nominal trials but the communication mode (4 levels) for each trial was randomly varied across the participants. Therefore, each flight crew saw all four off-nominal events but with each of the different communication modes such that across all flight crews there was a balanced presentation of off-nominal events paired with communication mode.

All 20 nominal trials were presented in the same order to the flight crews but was randomly ordered in which communication mode was paired with that trial. The nominal trials existed to provide the flight crews with experience with the various communication modes and the varying geometries of terminal maneuvering area procedures available at KORD. The nominal trials also served to increase the total number of trials to provide the necessary experimental matrix to introduce the off-nominal trials and to best replicate the operational conditions most likely to produce errors.

The dependent variables included quantitative and qualitative measures. The quantitative measures included 4DT adherence and conformance (e.g., commanded and actual speed delta), flight technical error (e.g., root-mean-squared-error of aircraft position from taxi centerline subtracting navigation system error), pilot control inputs (e.g., toe brake force, control wheel column and tiller inputs), off-nominal detection and response, and flight crew communication acts [42]. The qualitative measures included Likert-type post-run questionnaires, Situation Awareness Rating Technique [43], and NASA Task Load Index [44] and post-experiment paired comparisons for Situation Awareness (SA) – Subjective Workload Dominance (SWORD) [45].

Procedure

The flight crews were recruited from a NASA-established protocol. Each crew was given detailed pilot briefing materials prior to their arrival. All flight crews completed an informed consent and were provided a detailed briefing by the researchers upon their arrival. Participants were made aware of the experimental variables under test but were not cognizant that off-nominal events would be presented (this was confirmed in post-experimental debriefings). Three briefings in-total were presented over the two days of testing with each briefing focusing on the specific research objectives. A 90-minute training session was conducted prior to the start of each experimental phase, and pilots were trained to an asymptotic level of performance based on researcher observations and pilot feedback.

After training, twenty-four trials (20 nominal trials and 4 off-nominal trials) were conducted with the appropriate breaks included. Prior to each trial, the pilots were briefed on the evaluation task (e.g., approach to Runway 10) and were provided time to configure the aircraft for the operation (i.e., approach, departure, or taxi) and brief the procedure. Once the flight crew indicated they were ready, the trial began including a data-linked CPDLC message that was announced with a chime and displayed on the PFD CPDLC window and the EFB communication sub-menu. The F/O reviewed and acknowledged the message through the EFB (through a quick key access on main menu). Dependent upon the evaluation task, the trials lasted between 5 and 10 minutes. Pilots were asked to adhere to the speed and time commanded guidance to meet the RTA within a defined tolerance (+/- 5 kts and +/-30 sec), but to note any anomalies and to ensure safety of the aircraft at all times. After the trial ended, each pilot was asked to fill out several questionnaires that were administered on selected runs (12 experimental runs) that represented a cross-section of the nominal runs or were nominal analogs of the off-nominal trials.

RESULTS

The experiment, as a whole, had numerous objectives. The focus here is on the off-nominal results that reflect the efficacy of communications and display conditions to detect and recognize errors with flight-deck data-link CPDLC communications. Therefore, the off-nominal trials and the corresponding nominal trials data are presented here for comparison. The corresponding nominal trials were those that represented the routes and tasks presented during the off-nominal trials without the concomitant off-nominal error specific to each of the four off-nominal trials. In total, there were 8 nominal trials (i.e., 8 off-nominal trials across two experiments with 8 corresponding nominal trials presented during both experiments) and 4 off-nominal trials (along with 12 additional data runs) presented during this experiment. The 12 additional runs were intended to be "filler" runs to help provide enough trials to adequately mask the onset of the off-nominal data runs. Quantitative data was collected for these filler nominal trials and is presented elsewhere [46] but no qualitative data (i.e., post-run questionnaires) was gathered.

All significant results are presented using an *a priori* alpha (α) level of 0.05. Analyses of Variances (ANOVAs) were performed when appropriate Multivariate Analysis of Variance (MANOVA) tests were found to be significant. No significant results were found between Captain and FO qualitative ratings ($p > 0.05$) and, therefore, are grouped for analyses.

4DT Conformance and Taxi Performance

An ANOVA found no significant main effects across the four conditions for cross-track root-mean-squared error, $F(3,94) = 0.867$; ground speed deviation from commanded 4DT ground speed, $F(3,94) = 0.748$; and time deviation from RTA for the nominal analog trials, $F(3,94) = 1.180$. Generally, the flight crews ranged from 16.45 to 27.45 RMS error and maintained speeds within 1.5 knots of commanded groundspeed allowing the aircraft to conform to RTA times within 2.5 seconds of assigned RTA. Therefore, these results demonstrate that flight-crews are able to maintain ground speed commands and adhere to required-time-of-arrival clearances even under equivalent visual operational conditions. Table 1 presents the means for these dependent variables.

Table 1. Mean Cross-Track RMS, Ground Speed and RTA Deviations for Nominal Trials.

	Cross-Track RMS)	GS Deviation	RTA Deviation
Datalink + Voice	27.78 ft.	-0.031 kt.	-0.08 sec.
Datalink + Readback Only	16.45 ft.	-1.36 kt.	2.72 sec.
Datalink Only	22.34 ft.	1.1 7 kt.	-2.42 sec.
Datalink + Intent	21.69 ft.	-0.75 kt.	1.39 sec.

Off-Nominal Event Detection

Erroneous Data-link Clearance. There were 3 off-nominal events per experiment condition across the 12 flight crews . Ten of the flight crews correctly recognized the erroneous data-link clearance (the inclusion of "Mike Four" in the taxi clearance) and contacted ATC to query the clearance and/or reject the clearance. The two instances where the erroneous clearance was not detected occurred for the data-link only condition.

In the ten correct identifications, there was an observed difference in how the error was identified with flight crews. In either of the two voice conditions (data-link + voice; data-link + read-back only), the crews immediately recognized that the CDPLC message did not correspond with the voice clearance. In contrast, the data-link + intent flight crews relied on review of the airport diagram and surface map display. This review took significantly longer to detect the error and contact ATC. These three conditions, however, were much better in detection of the off-nominal situation than the data-link only condition in which only one flight crew successfully rejected the clearance.

In the other two off-nominal scenarios (66% of the data-link only condition) accepted the clearance and followed the taxi instructions including turning back toward the active runway onto Mike Four at ORD resulting in the aircraft being stuck at the hold-short line for Runway 10 and the need to contact the tower for permission to cross the active runway.

The data supported our hypothesis that voice (as redundant to data-link) would result in greater awareness and likelihood of detection of the error. Explicit display of intent information could also be successfully used to mitigate these off-nominal conditions but it does not appear to be as efficient.

Data-Link Clearance Received for Other Aircraft. All the flight crews with either the data-link only or data-link + intent condition failed to detect the data-link clearance received intended for another aircraft and accepted the new clearance despite being 1 nm on final landing on Runway 22R. Upon taxi, these flight crews realized that the taxi instructions resulted in a route disconnect from the active runway to the ramp area and contacted ATC to request new taxi instructions. However, none of the six flight crews with data-link + voice or data-link + read-back accepted the data-link clearance and correctly recognized that the data-link message was not specific to their aircraft. These flight crews ignored the CDPLC message and contacted ATC to query the message after turn-off.

These data supported our hypothesis that the flight crew would not properly attend to the CPDLC message and would accept the clearance without thorough review. This contrasts with the flight crews who also heard the voice clearance, and properly recognized that the taxi clearance was not congruent with the landing runway and was not intended for their aircraft.

Taxiway Conflict. The objective of the scenario was to determine if any of the experimental conditions would result in significant head-down time and increased intra-cockpit communications that may prevent the timely identification of the other traffic. (The scenario was not intended to result in a collision but would have resulted in a nose-to-nose situation requiring ground vehicles to tow the aircraft if not detected in an adequate amount of time). The data showed that none of the flight crews – irrespective of the experimental condition - experienced a taxiway conflict. Because all the conditions had a HUD with a forward-looking-infrared (FLIR) enhanced vision system, the Captain was able to detect the traffic taxiing toward them on the same taxiway. Our hypothesis was that the flight crew that had either of the voice and data-link conditions would more easily recognize the potential conflict and would be proactive in avoiding the situation. This result was neither support nor refuted. Instead, the results showed the significant advantage that an enhanced vision system provides for low-visibility operations, especially in this scenario.

Approach Runway Incursion. For the data-link only condition, there was one runway incursion (33% of the flight crews) that occurred because the flight crew was entirely unaware that an aircraft was in position-and-hold awaiting take-off clearance when the aircraft landed on the active runway. For the data-link + intent condition, the flight crews were able to adequately recognize and avoid the incursion but were slow in identifying the evolving situation and reported significant workload in interfacing with the traffic intent data and determining what the aircraft was intending to do. In contrast, the flight crews in both the voice

conditions began discussing the aircraft on the active runway upon hearing the clearance and then recognizing that the incursion aircraft had not taken-off and had ample time to discuss options, initiate a go-around, and contact ATC as to their situation.

These data supported our hypothesis that having party-line information would cue the flight crew that an aircraft was holding on the landing runway. Also, having data-link with traffic intent might also alert the flight crew to the runway hazard. Two runway incursions were *averted* in the data-link only condition, when in fact, our hypothesis would suggest this would not be the case. The use of an enhanced vision-equipped HUD allowed the Captain to see the incurring traffic on the runway and execute a go-around.

Situation Awareness

Situation Awareness Rating Technique (SART). In Table 2, the SART data for qualitative post-run questionnaire items for off-nominal trials by experimental condition are presented. The SART uses the constructs of: 1) demand on attentional resources; 2) supply of attentional resources; and, 3) understanding. Mean values for SART were computed and are shown in Table 2, where the SART rating is "understanding" reduced by the difference of "demand" minus "supply" (i.e., SART = [understanding – (demand – supply)]).

ANOVAs on the mean SART ratings found significant main effects across the four experimental conditions for all four off-nominal events.

- For the "erroneous data-link clearance" event, there was a significant effect found, $F(3,24) = 5.529$. A Student-Newman-Keuls (SNK) post-hoc test showed that pilots rated the data-link + voice and data-link + read-back conditions to be significantly higher than the data-link only and data-link only + intent conditions.
- The same significant differences in SART ratings by condition was found for the "approach runway incursion" event, $F(3,24) = 7.682$.
- However, SNK post-hoc tests on the significant main effects for both the "taxiway conflict" ($F(3,24) = 13.496$), and "data-link clearance received for other aircraft" ($F(3,24) = 7.362$) events, evinced that the data-link + read-back condition was rated significantly higher for situation awareness compared to the other three experimental conditions.

Table 2. Mean Pilot Responses for Qualitative Post-Run Questionnaire Items - Off-Nominal Event

| Experimental Condition	Off-Nominal Event	Q1	Q2	Q3	Q4	Q5	Q6	Q7	SART	NASA-TLX
Data-Link + Voice	Erroneous Data-link Clearance	1.50	1.33	1.33	2.00	1.67	1.33	1.17	5.83	26.53
Data-Link + Readback	Erroneous Data-link Clearance	1.67	1.50	2.00	2.83	2.33	1.83	2.33	6.17	26.81
Data-Link Only	Erroneous Data-link Clearance	1.83	2.33	2.50	2.83	3.00	2.33	2.50	6.33	28.47
Data-Link + Intent	Erroneous Data-link Clearance	2.33	3.00	2.83	3.33	3.00	2.83	2.67	10.00	32.78
Data-Link + Voice	Data-Link Clearance Received for Other Aircraft	1.50	2.33	2.00	2.17	2.50	2.17	2.00	0.00	42.50
Data-Link + Readback	Data-Link Clearance Received for Other Aircraft	2.00	2.50	2.33	3.00	2.67	2.50	2.50	2.66	47.92
Data-Link Only	Data-Link Clearance Received for Other Aircraft	2.00	3.33	3.33	3.33	3.67	3.17	3.33	3.66	50.56
Data-Link + Intent	Data-Link Clearance Received for Other Aircraft	3.17	4.50	3.50	3.50	3.83	3.33	3.67	10.67	51.80
Data-Link + Voice	Taxiway Conflict	1.17	1.00	1.17	2.50	1.00	1.50	1.00	1.16	26.94
Data-Link + Readback	Taxiway Conflict	1.50	1.50	1.50	2.83	3.33	1.83	3.17	3.50	31.94
Data-Link Only	Taxiway Conflict	2.00	2.17	2.00	3.00	3.83	2.17	3.67	4.67	36.11
Data-Link + Intent	Taxiway Conflict	2.00	2.33	2.33	3.67	4.33	2.33	3.67	10.33	40.56
Data-Link + Voice	Approach Runway Incursion	1.00	1.00	1.17	1.83	1.50	1.50	1.83	5.50	29.86
Data-Link + Readback	Approach Runway Incursion	1.17	1.00	1.17	2.50	2.83	1.67	1.83	5.50	20.69
Data-Link Only	Approach Runway Incursion	1.17	1.67	1.67	2.50	3.50	2.00	2.17	9.17	38.89
Data-Link + Intent	Approach Runway Incursion	1.83	2.00	1.83	4.33	4.17	3.33	2.67	10.83	43.75

In comparison, the subjective SA data for the corresponding nominal trials are presented in Table 3. These data show a significant main effect for SART, F(3,95) = 40.675. A SNK post-hoc test revealed three subsets: (a) data-link-only (1.54), (b) data-link + intent (3.46) and data-link + voice (3.92) and (c) data-link + read-back (10.25). Therefore, pilots rated the read-back only condition to be significantly higher for situation awareness than the three other conditions and the data-link only to be the worse concept for situation awareness during the nominal trials.

Table 3. Mean Pilot Responses for Qualitative Post-Run Questionnaire Items - Nominal Trials

Experimental Condition	Q1	Q2	Q3	Q4	Q5	Q6	Q7	SART	NASA-TLX
Data-Link + Voice	1.83	1.71	1.46	2.54	2.83	2.00	2.21	3.92	43.31
Data-Link + Readback	1.38	1.21	1.54	1.58	1.21	1.50	1.33	10.25	23.59
Data-Link Only	3.17	3.33	3.25	3.67	4.50	2.96	4.29	1.54	54.07
Data-Link + Intent	1.83	1.61	1.92	2.54	3.21	1.92	2.83	2.46	35.32

Situation Awareness - Workload Dominance (SA-SWORD) Scale. In Figure 5, the geometric-means for the paired comparison scales for each experimental condition are shown. As a post-test, pair-comparison test, the rating data does not differentiate between nominal and off-nominal trials.

An ANOVA revealed a significant main effect for display condition for SA-SWORD, F(3, 69) = 40.601. Post-hoc test showed that pilots rated the data-link + read-back condition to be significantly higher for situation awareness than the other three conditions. No other effects were found to be significant.
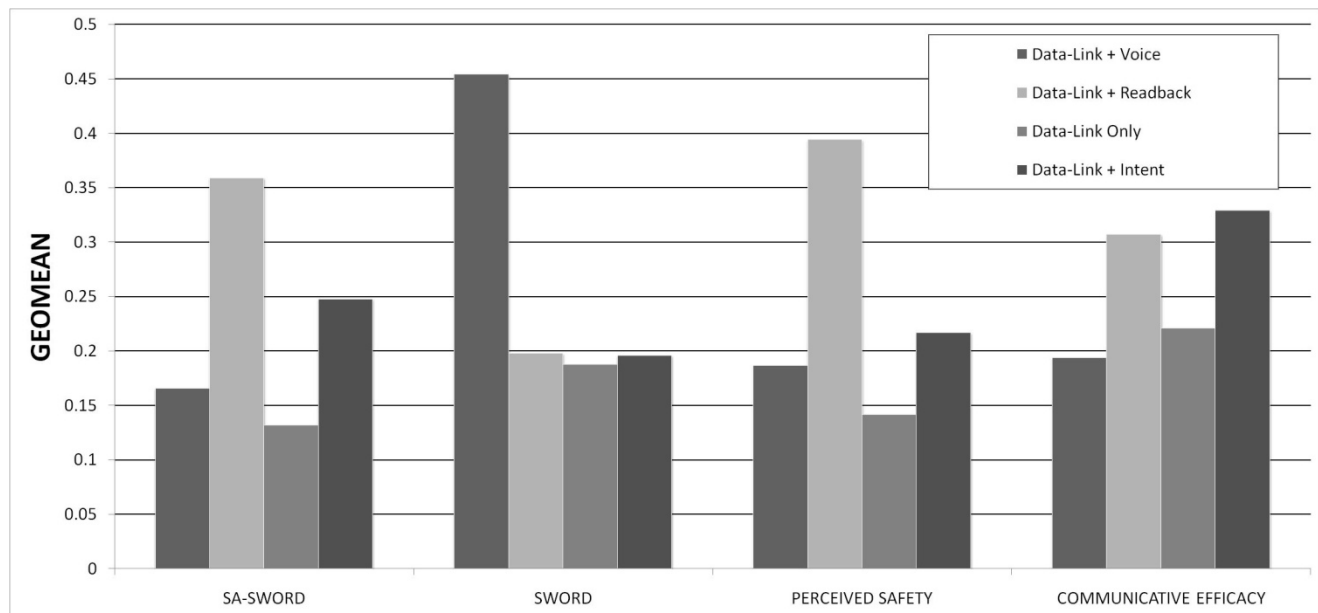
Figure 5.   Geometric-Mean Values for Paired Comparison Scales

Mental Workload

NASA Task-Load-Index (NASA-TLX). The post-run work was assessed using the NASA TLX scale. The mean NASA-TLX workload for the off-nominal trials are shown in Table 2 and the corresponding nominal trial data are shown in Table 3, broken down by experimental condition. An ANOVA found no significant main effect for NASA-TLX for any of the four off-nominal trials, $p > 0.05$. The four experimental conditions were found to be equal in perceived mental workload as rated by the NASA-TLX scale.   However, for the nominal trials, there was a significant main effect for NASA-TLX, $F(3,95) = 11.719$. SNK post-hoc tests revealed three subsets where the pilots rated the data-link + read-back condition (23.58) to be significantly lower for workload than the other three display conditions. The data-link + intent (35.32) and data-link + voice (43.31) were both rated significantly lower than data-link (54.07).

Subjective Workload Dominance (SWORD) Scale.   In Figure 5, the geometric-means for the paired comparison SWORD data are shown. As a post-test, pair-comparison test, the rating data does not differentiate between nominal and off-nominal trials. An ANOVA revealed a significant main effect for display condition for SWORD, $F(3,69) = 68.169$. Post-hoc comparisons evinced that pilots rated the data-link + voice (data-link as redundant to voice) significantly higher in mental workload on the SWORD post-experiment paired-comparison scale compared to the other three conditions which did not differ statistically from each other ($p > 0.05$).

Post-Run Subjective Assessments

Following the runs, the flight crews rated seven questions using a 5-point Likert-scale (Figure 6). The mean ratings to each of these questions are given for the off-nominal trials in Table 2 and the corresponding nominal trial data are shown in Table 3, broken down by experimental condition. Only the statistically-significant results are reported herein.

Post-Run Ratings	Excellent	Good	Acceptable	Almost Acceptable	Absolutely Unacceptable
Please rate your level of situation awareness for:	1	2	3	4	5
Q1. Overall navigation awareness (e.g., direction of travel)					
Q2. Taxi route awareness during final/roll-out (e.g., awareness of taxi route during approach and landing)					
Q3. Taxi route awareness during taxiing					
Q4. Awareness of other aircraft position					
Q5. Awareness of other aircraft clearances and "partyline"					
Q6. Awareness of time to complete task (e.g., time remaining to complete 4DT clearance; required-time-of-arrival constraints)					
Q7. Please provide your overall rating of perceived safety taxiing aircraft during the trial					

Figure 6. Post-Run Likert Questionnaire [34]

Navigation Awareness (Q1). Pilots did not report, for the post-run questionnaire, any statistically significant differences in perceived navigation awareness for any of the display experimental conditions across the four off-nominal events, $p > 0.05$. There was, however, a significant main effect reported for the nominal trials, $F(3,95) = 18.99$. Pilots rated the data-link only to provide significantly less navigation awareness than the other three conditions.

Taxi Route Awareness (Q2, Q3). An ANOVA revealed a significant main effect for the "data-link clearance received for other aircraft" off-nominal event, $F(3, 24) = 5.193$. A SNK post-hoc test found two subsets: (a) the flight crews reported that the data-link + intent was rated significantly higher for taxi route awareness for all conditions except data-link only and (b) data-link + voice was rated not significantly different than either the data-link + read-back and data-link only conditions ($p > 0.05$).

For the nominal trials, an ANOVA revealed a significant main effect for taxi route awareness, $F(3, 95) = 28.47$. A SNK post-hoc test showed that pilots rated the data-link only condition to be significantly lower for taxi route awareness than the other three conditions.

Traffic Awareness (Q4). ANOVAs performed on the post-run questionnaire item for "traffic awareness" found a main effect for the "taxiway conflict" off-nominal event only, $F(3,24) = 112.789$. A SNK post-hoc test evinced that pilots rated that: (a) data-link + read-back condition to be significantly higher for traffic awareness than data-link only condition but not the data-link + voice or data-link + intent experimental conditions, and, (b) there was no significant differences between data-link only and data-link + voice and data-link + intent conditions. For the nominal trials, the same pattern of results emerged, $F(3,95) = 17.214$, with two SNK post-hoc subsets identical to the off-nominal results for qualitative traffic awareness ratings.

4DT Awareness (Q6). No significant differences were found for qualitative differences in pilot post-run questionnaire response to "awareness to complete task (e.g., time remaining to complete 4DT clearance; required-time-of-arrival constraints)", $p > 0.05$. For the nominal trials, there was a significant main effect, $F(3,95) = 8.49$. Pilots rated the data-link only condition to be significantly lower for 4DT awareness than the other three display conditions which did not differ statistically from each other.

Communication Awareness

Post-Run Questionnaire (Q5). Significant main effects were found for display experimental condition for post-run questionnaire item for communication awareness for the off-nominal events of "taxiway conflict", $F(3,24) = 7.527$, and "approach runway incursion, $F(3,24) = 8.842$. SNK post-hoc tests revealed that the data-link + read-back condition was rated significantly higher for data-link + read-back only than the other display conditions which did not differ for communication awareness.

Analyses for qualitative reports of communication awareness taken during the nominal trials also revealed a significant main effect, $F(3,95) = 44.49$. SNK post-hoc test showed three subsets: (a) data-link only was rated lowest for awareness of aircraft clearances and party-line information, (b) data-link + intent and data-link + voice was rated higher than data-ink only but poorer for communication awareness compared to data-link + read-back, and (c) data-link + read-back was qualitatively rated the highest for the construct than the other three display conditions.

Post-Experimental Paired Comparison. An ANOVA revealed a significant main effect for display condition for communication efficacy, $F(3,69) = 29.635$. Post-hoc LSD comparisons showed that pilots rated the data-link + voice and data-link only conditions to be significantly lower for communication effectiveness both compared to data-link + Intent and data-link + readback only conditions.

Flight Crew Communication Acts. An observational analysis was conducted on video taken during the off-nominal trial to evaluate crew coordination based on communication acts because past research has reported that flight crews increase the number of speech acts with data-link compared to voice communications. The data-link only and data-link + intent conditions had an increased number of communications between Captain and FO. Based on Kanki, Lozito, and Foushee [34] Table 4 presents the overall initiating speech ratio (I) and overall response speech ratio (R) structured by commands (COM), questions (QST), observations (OBS), and dysfluencies (DYS).

Perceived Safety

Post-Run Questionnaire (Q7). Results found significant main effects for pilot perceptions of perceived safety on the post-run questionnaire during the off-nominal event of "erroneous data-link clearance", $F(3,24) = 5.501$. SNK post-hoc test showed that pilots ranked the data-link + read-back condition as significantly higher in overall rating of perceived safety of the aircraft during that event. No other experimental conditions were significantly different from each other.

During the nominal trials, however, the main effect for perceived safety during taxi, $F(3,95) = 27.29$, showed three post-hoc SNK subset: (a) data-link only was rated the lowest for perceived safety, (b) followed by data-link + intent and data-link + voice, and (c) data-link + read-back which was rated the highest for perceived taxi safety during the nominal trials.

Post-Experimental Paired Comparison. An ANOVA revealed a significant main effect for display condition for SWORD, $F(3,69) = 30.782$. Post-hoc LSD comparisons evinced that pilots rated the data-link + readback only to be significantly higher for pilot perceptions of safety for surface operations compared to the other three display conditions. The data-link only condition was rated significantly lowest for perceived safety than the other three display conditions.

Table 4. Mean I x R Communication Acts during Off-Nominal Trials

		Captain → FO		
	Datalink Only	Datalink + Intent	Datalink + Voice	Datalink + Readback
COM	8	6	3	4
QYS	15	14	7	3
OBS	21	18	4	6
DYS	6	5	2	2
		FO → Captain		
	Datalink Only	Datalink + Intent	Datalink + Voice	Datalink + Readback
COM	N/A	N/A	N/A	N/A
QYS	14	17	4	2
OBS	12	16	3	5
DYS	5	3	0	1

DISCUSSION

The research objective was to evaluate the flight deck impacts of different data-link implementations that may be considered as options for ATC-pilot communications for NextGen. The purpose of the research was to

specifically evaluate a voice-by-exception approach within a NextGen terminal maneuvering area operational concept and also, to explore other possible approaches that may ameliorate issues reported in past voice-by-exception studies while providing reductions in voice frequency congestion. To accomplish the stated objective, four conditions were examined under potential NextGen operational conditions. The four conditions were data-link only (voice-by-exception), voice + data-link (all voice clearances and read-backs supplemented by CPDLC), voice + read-back only (data-link ATC clearances and mandatory pilot read-back of CPDLC clearance over voice channel), and data-link + intent (data-link only with the addition of other traffic intent data shown on the navigation display). Off-nominal testing results were presented, with nominal trial data for comparison, to evaluate the performance and safety of these communications modalities in the event of errors or blunders. Four off-nominal scenarios were specifically designed to evaluate the efficacy of these different communications concepts.

Across all off-nominal conditions, the results demonstrated that data-link only is not an optimal solution for these operations and corroborates previous research which advises caution for a voice-by-exception system because of the significant concerns for reduction of party-line and situation awareness on the part of the flight crew. The voice-by-exception experimental condition was almost universally the highest workload and lowest situation awareness condition.

Our belief that providing traffic intent data to the flight crew might offset the loss of party-line awareness in the data-link only condition was not fully supported. Although the pilots reported the intent data was helpful - providing increases in traffic awareness and situation awareness - it required significant cognitive overhead to interpret the data and utilize the information to develop a mental model of what the other traffic were doing without the benefit of the party-line chatter. Intent information provided some aspects of increased "party-line" awareness, but was not universally effective as witnessed by the "Data-Link Clearance Received for Other Aircraft" off-nominal where this condition was just as deficient as the data-link only case.

As expected, the data-link + voice condition was well-received by the flight crew because it represented a condition they were familiar with, yet it also supplemented the voice clearances with textual data-link messages on the EFB and PFD CPDLC message area. However, the flight crew reported significantly increased workload and concerns regarding the rapid pace of voice communications that is typical of airports such as Chicago O'Hare. This condition would not reduce voice congestion, and it would almost certainly be unacceptable to controllers as they would always be sending two messages (voice and data-link), instead of just one.

The subjective and off-nominal data showed that the data-link + read-back condition was clearly the best option across almost all the dependent variables tested. By requiring pilot read-back of CPDLC, flight crew responded that they were able to retain the party-line broadcast awareness because they knew what the other traffic was "intending" to do while at the same time reducing the voice congestion by almost 50%. This also significantly freed up the frequency to allow the flight crews to more easily contact the controller if necessary (e.g., non-routine communications). The data-link CPDLC was welcomed because it allowed the pilots the time allowance to review and accept the clearance and response as time permitted because of the reduction of time pressure (which was mentioned as a significant stressor by almost the all flight crews with experience at KORD) that is often present at major U.S. airports.

The off-nominal results confirmed the finding that voice for routine communications significantly enhances situation awareness for terminal maneuvering area operations. During the approach runway incursion scenario, the flight crews recognized the situation through the party-line (incursion aircraft read-back of clearance) well before the data-link + intent flight crews identified the traffic on the surface map as presenting a potential hazard. For the data-link only condition, there was a runway incursion that occurred because the

flight crew was entirely unaware that an aircraft was in position-and-hold awaiting take-off clearance when the aircraft landed on the active runway. Similar results were found for the other off-nominal events (except the taxiway conflict) in which the data-link only condition significantly reduced the flight crews ability to detect impending hazardous situations and either failed to detect the error and/or other traffic or were highly reactive in executing avoidance maneuvers. Overall, the results demonstrate the substantial importance of providing some party-line capability for any new ATC-pilot data communication system. The conclusion drawn here is that a promising solution may be the use of mandatory pilot read-back of CPDLC data-link routine clearances in the terminal maneuvering area and the use of voice communication for all non-routine and exigent communications. The conclusion matches that observed during the CASCADE D-Taxi trials at Brussels [47] which recommended the use of mandatory pilot read-back as, "…effective mitigation means against erroneous or non-intended clearance details in the sent message" (p. 83).

FUTURE DIRECTIONS

Although the issues of data-link only have been well-documented in the literature for aircraft operations typical of today's system, this "voice-by-exception" data-link protocol continues to remain the primary communications end-state. The conclusions drawn from the results presented herein mirror other research and share the concerns for implementation of a safe and effective data-link only system for the terminal maneuvering area. Although voice-by-exception permits voice communication - and requires voice for non-routine communications - research suggests that there are ample situations and opportunities in which a data-link only approach may reduce safety margins. The use of mandatory pilot read-back of CPDLC data-link routine clearances in the terminal maneuvering area will be explored further.

Research will continue in the investigation on the effects of data-link communications for NextGen 4DT surface operations. Continued exploration of methods and means for graphical display of traffic intent will continue to possibly mitigate, or eliminate, anticipated effects of voice-by-exception flight-deck communications. To further enhance the safety of these technologies, NASA is also investigating conflict detection and resolution (CD&R) technologies to provide an additional protective safety layer in the event that the tactical or strategic awareness of the flight crew is compromised and a potentially hazardous situation is presented. The system integrates airborne and ground-based technologies which include flight deck displays, conflict detection and alerting algorithms, on-board position determination systems, airport surveillance systems, and controller-pilot data link communications. Finally, the present results demonstrate that flight crews were able to adhere to required-time-of-arrival clearances without significant effort even under equivalent visual operational speeds and traffic densities. Future research will focus on optimization of 4DT algorithms and flight methodologies to meet these and other challenges to NextGen operations.

REFERENCES

[1] Prinzo, O.V. (1996). An Analysis of Approach Control/Pilot Voice Communications. FAA Civil Aeromedical Institute, (DOT/FAA/AM-96/26), p. 3-44.

[2] Cardosi, K.M. (1993): An Analysis Of En Route Controller-Pilot Communications. Report No. DOT/FAA/RD - 93/11. Washington, DC: Department of Transportation, Federal Aviation Administration.

[3] Morrow, D and Rodvold, M. (1993) The Influence of ATC Message Length and Timing on Pilot Communication. NASA Contractor Report 177621. August 1993

[4] Cushing, S. (1994). Fatal Words. Chicago, IL: University of Chicago Press.

[5] Eurocontrol (2005): Initial Communications Operating Concepts And Requirements For The Future Radio System. Eurocontrol/FAA Future Communications Study Operational Concepts and Requirements, January 2005.

[6] Kerns, K. (1991): Data-Link Communication Between Controllers and Pilots: A Review and Synthesis of the Simulation Literature. International Journal of Aviation Psychology 1991, Vol. 1, No. 3, Pages 181-204.

[7] Federal Aviation Administration (1995): User Benefits of Two-Way Data Link ATC Communications: Aircraft Delay and Flight Efficiency in Congested En Route Airspace. Federal Aviation Administration Data Link Benefits Study Team FAA Report FAA/CT-95-4, February 1995.

[8] Federal Aviation Administration (1996a): The Interfaces Between Flightcrews and Modern Flight Deck Systems. Federal Aviation Administration Human Factors Team Report. June 1996.

[9] Federal Aviation Administration (1996b). Benefits of Controller-Pilot Data Link ATC Communications in Terminal Airspace. Federal Aviation Administration Data Link Benefits Study Team. FAA Report FAA/CT-96-3, September 1996.

[10] Lozito, S. A., McGann, S. A. and Corker, K. (1993): Data link air traffic control and flight deck environments: Experiments in flight crew performance. In R. E. Jensen and D. Neumeister (Eds.), Proceedings of the 7th International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University, 1993, pp. 1009-1015.

[11] Corwin, W.H. and McCauley, H. (1990): Considerations for the retrofit of data link. SAE, Aerospace Technology Conference and Exposition, Long Beach, CA; United States; 1-4 Oct. 1990.

[12] Andre, A.D., Lins, J.M.C., and Wilson, J. (2003): Conveying Message Criticality Via Datalink. Proceedings of the Twelfth International Symposium on Aviation Psychology, pp. 54-59. Dayton, OH: Wright State University.

[13] Talotta, N. J., and Shingledecker, C. (1992): Controller Evaluation Of Initial Data Link Terminal Air Traffic Control Services: Mini-Study 3, Report No. DOT/FAA/CT-92/18, Volume I. Washington, DC: U.S. Department of Transportation, Federal Aviation Administration. 1992.

[14] Groce, J. L. and Boucek, G. P. (1987): Air transport crew tasking in an ATC data link environment. SAE Tech. Paper 871764. Warrendale, PA: SAE International. 1987

[15] van Gent, R.N.H.W. (1995): Human Factors Issues With Airborne Data Link; Towards Increased Crew Acceptance For Both En-Route And Terminal Flight Operations. National Aerospace Laboratory NLR TP 95666, Amsterdam, NL, 1995.

[16] Waller, M. C., & Lohr, G. W. (1989): A Piloted Simulation Of Data Link ATC Message Exchange. NASATech. Paper 2859. Hampton, VA: NASA Langley Research Center. 1989.

[17] Knox, C. E. and Scalon, C. H. (1991): Flight Tests With A Data Link Used For Air Traffic Control Information Exchange. NASA Tech. Paper 3135. Hampton, VA: NASA Langley Research Center, 1991

[18] Midkiff, A. H., and Hansman, R. J., Jr. (1993): Identification Of Important "Party Line" Information Elements And Implications For Situational Awareness In The Datalink Environment. Air Traffic Control Quarterly, 1(1), 5-30.

[19] Navarro, C. and Sikorski, S. (1999): Datalink Communication in Flight Deck Operations: A Synthesis of Recent Studies. International Journal of Aviation Psychology 1999, Vol. 9, No. 4, Pages 361-376.

[20] Williams, J., Hooey, B.L., & Foyle, D.C. (2007). Pilot conformance to time-based taxi clearances: Implications for advanced surface traffic management systems. Human Centered System Lab Technical Report (HCSL-07-02). December, 2007.

[21] D. C. Foyle, A. D. Andre, R. S. McCann, E. Wenzel, D. Begault, and V. Battiste. Taxiway navigation and situation awareness (T-NASA) system: Problem, design philosophy and description of an integrated display suite for low-visibility airport surface operations. SAE Transactions: Journal of Aerospace, 105:1411{1418, 1996.

[22] R. S. McCann, B. L. Hooey, B. Parke, D. C. Foyle, A. D. Andre, & B. Kanki. An evaluation of the taxiway navigation and situation awareness (T-NASA) system in high-fidelity simulation. SAE Transactions: Journal of Aerospace, 107:1612{1625, 1998.

[23] Jones, D. R., Quach, C. C., and Young, S. D. (2001): Runway Incursion Prevention System – Demonstration and Testing at the Dallas/Fort Worth International Airport, Proceedings of the 20th Digital Avionics Systems Conference, 2001.

[24] Cheng, V.H. L., Yeh, A., Diaz, G.M., & Foyle, D.C. (2004). Surface-operation benefits of a collaborative automation concept. Proceedings of the AIAA Guidance, Navigation, and Control Conference, Paper AIAA 2004-5409.

[25 Rathinam, S., Montoya, J., and Jung, Y. (2008). An optimization model for reducing aircraft taxi times at the Dallas Fort Worth International Airport. 26th International Congress of the Aeronautical Sciences (ICAS) Anchorage, Alaska, Sep. 14-19, 2008.

[26] Arthur, J.J., Prinzel, L.J., Bailey, R.E., Shelton, K.J., Williams, S.P., Kramer, L.J., Norman, R.M. (2008). Head-Worn Display Concepts for Surface Operations for Commercial Aircraft. NASA TP-2008-215321

[27] National Transportation Safety Board (2008): NTSB Most Wanted List Transportation Safety Improvements, www.ntsb.gov/recs/mostwanted/index.htm, November 2008.

[28] Federal Aviation Administration (2008) FAA Runway Safety Report. Federal Aviation Administration Office of Runway Safety, June 2008.

[29] National Transportation Safety Board (2000) Safety Recommendation, Letter to the FAA Administrator, A-00-66, July 2000.

[30] Federal Aviation Administration (2000): Introduction to TCAS II, Version 7.

[31] Ivanescu, D., Powell, D., Shaw, C., Hoffman, E., and Zeghal, K., (2004): Effect of Aircraft Self-Merging in Sequence on an Airborne Collision Avoidance System, AIAA Guidance, Navigation, and Control Conference and Exhibit, AIAA 2004-4994.

[32] Pritchett, A.R. and Hansman, R.J. (1995): Variations In "Party Line" Information Importance Between Pilots Of Different Characteristics. Presented At The Eighth International Symposium On Aviation Psychology Columbus, Oh, April, 1995.

[33] RTCA (2008): Two New Committees Approved at PMC: Terrain and Airport Databases and Future ADS-B/TCAS Relationships, RTCA Digest, No. 182.

[34] Foyle, D.C., Williams, J.L., & Hooey, B.L. (2008). Toward the development of information requirements of time-based (4-D) taxi clearances: Pilot-in-the-loop results. NASA Airspace Systems Program Technical Interchange Meeting, Austin, TX March 17-20, 2008.

[35] Joint Planning & Development Office (2007): Concept of Operations for the Next Generation Air Transportation System, Version 2.0, June 2007.

[36] Green, D. F. (2006): Runway Safety Monitor Algorithm for Single and Crossing Runway Incursion Detection and Alerting, NASA/CR-2006-214275.

[37] Jones, D. R. (2002): Runway Incursion Prevention System Simulation Evaluation. Proceedings of the 21st Digital Avionics Systems Conference, Irvine, CA, Oct. 27 – 31, 2002

[38] Jones, D. R. (2005): Runway Incursion Prevention System Testing at the Wallops Flight Facility, Proceedings of the SPIE Defense & Security Symposium, 2005.

[39] D. R. Jones and J. M. Rankin. A system for preventing runway incursions. Journal of Air Traffic Control, 44(3), July { September 2002

[40] Jones, D. R., and Prinzel, L. J., III (2006): Runway Incursion Prevention for General Aviation Operations, Proceedings of the 25th Digital Avionics Systems Conference, 2006.

[41] Joint Planning & Development Office (2007): Concept of Operations for the Next Generation Air Transportation System, Version 2.0, August, 2007)

[42] Kanki, B., Lozito, S., & Foushee, H.C. (1989). Communication indices of crew coordination. Aviation Space and Environmental Medicine, 60, 56-60.

[43] Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. AGARD Conference Proceedings No 478, Situational Awareness in Aerospace Operations. Aerospace Medical Panel Symposium, Copenhagen. October. AGARD-CP-478. pp 3-1 - 3-37.

[44] Hart, S. G. & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. Human mental workload.

[45] Vidulich, M. A. and Hughes, E. R. (1991). Testing a subjective metric of situation awareness. In Human Factors Society 35th Annual Meeting. pages 1307-1311. Santa Monica, CA. Human Factors Society.

[46] Shelton, K., Prinzel, L.J., Arthur, J.J., Jones, D.R., Allamandola, A.S., & Bailey, R.E. (2009). Data-link and surface map traffic intent displays for NextGen 4DT and Equivalent Visual Surface Operations. Proceedings of the SPIE, Volume 7328, pp. 73280C-73280C-18

[47] Stephenson, K. (2007). D-TAXI trial project (DTP) final report. Brussels, Belgium: EUROCONTROL CASCADE.

ACKNOWLEDGMENTS