Bakowski, D.L., Hooey, B.L. and Foyle, D.C. (2017). Flight deck surface trajectory-based operations (STBO): A four-dimensional trajectory (4DT) simulation. Proceedings of the 36th Digital Avionics Systems Conference (DASC 2017), St. Petersburg, FL, Sept. 2017.

Flight Deck Surface Trajectory-Based Operations (STBO): A Four-Dimensional Trajectory (4DT) Simulation

Deborah L. Bakowski
San José State University
NASA Ames Research Center
Moffett Field, CA, USA
debi.bakowski@nasa.gov

Becky L. Hooey
NASA Ames Research Center
Moffett Field, CA, USA
becky.l.hooey@nasa.gov

David C. Foyle
NASA Ames Research Center
Moffett Field, CA, USA
david.foyle@nasa.gov

Abstract— **In four-dimensional trajectory (4DT) Surface Trajectory-Based Operations (STBO), aircraft are assigned a conflict-free 4DT which defines an expected location (*x,y coordinates*) at all times, *t*, along the taxi route (with *altitude*, being fixed). These 4DTs afford the highest temporal certainty at *all* points along the taxi route, and at the departure runway. In the present study, a 4DT flight deck display was presented on the Airport Moving Map (AMM) to support pilot conformance to a 4DT clearance while taxiing under manual control. This pilot-in-the-loop simulation compared the effect of 4DT flight deck display formats on distance from the expected 4DT location, conformance to the displayed tolerance band, eyes-out time, and pilot ratings of safety and workload. In the *defined-tolerance* display format, a graphical representation of the expected 4DT location, with a distance-based allowable-tolerance band, was depicted on the AMM. Two defined-tolerance band sizes were tested +/-164 ft and +/-405 ft. In the *undefined-tolerance* display format, the expected 4DT location was displayed graphically on the AMM, with no indicated allowable-tolerance bounds. Each taxi trial included 4DT speed changes (two or five, per trial) and a range of 4DT taxi speeds. Results showed that the larger (+/-405 ft) defined-tolerance band yielded higher conformance levels than the smaller (+/-164 ft) band, with pilots staying within the specified and displayed conformance bounds more in the larger (99.71%) than the smaller defined-tolerance band (93.37%). However, in terms of being able to predict the location of the aircraft compared to the expected 4DT location, the smaller defined-tolerance band resulted in pilots keeping their aircraft closer to the 4DT location, for both average distance and for a given confidence interval (e.g., 95%), than either the larger defined-tolerance band or the undefined-tolerance display format. The larger tolerance band yielded more "eyes out-the-window" time than the smaller tolerance band. Pilots also rated taxing with the larger tolerance band as safer than the smaller tolerance band.**

Keywords—STBO, 4DT, flight deck, displays, taxi, surface ops

I. INTRODUCTION

The Next Generation Air Transportation System (NextGen) [1] envisions improving the safety and efficiency of airspace operations, while reducing the environmental impacts and increasing capacity. To realize these benefits, NextGen concepts integrate new technology, automation, and procedures into all phases of flight, including surface operations.

In current-day surface operations, aircraft are generally handled on a first-come, first-served basis which contributes to: uncertainty about taxi durations and when aircraft will arrive at their departure runway, causing congestion and long departure-queues [2]; stop-and-go taxi which contributes to excess fuel consumption [3]; and, an inability to support other NextGen concepts that depend on precise departure times [4].

A number of NextGen concepts introduce information sharing and scheduling/sequencing management tools into gate, surface, runway, and terminal-area operations to support more efficient surface operations and increased throughput. For example, the integrated arrival, departure, and surface (IADS) concept creates an integrated schedule of arrivals and departures and uses information sharing to support scheduling [5], and the Surface Collaborative Decision Making (S-CDM) concept provides users with access to aircraft surface surveillance data and flight status information [6]. In Europe, Airport Collaborative Decision Making (A-CDM) provides users with access to more accurate and higher-quality information, particularly in the pre-departure phase [7].

Surface Trajectory-Based Operations (STBO) is a concept of operations for managing flows and resources on the airport surface to improve the efficiency, throughput, and predictability of surface operations while reducing the environmental impact [8]. STBO envisions delivering a specific aircraft to a specific place on the airport (e.g., runway) at a specific time to meet a specific event (e.g., takeoff) in the most efficient manner possible [8].

II. SURFACE TRAJECTORY-BASED OPERATIONS (STBO)

A. NextGen STBO Operations

In the NextGen timeframe, the STBO concept is expected to provide Air Traffic Control (ATC) with decision support tools (DSTs) to support capabilities like departure runway scheduling and sequencing. DSTs will provide sequencing suggestions for aircraft in the runway queue and time-based recommendations for traffic management [9]. Similarly, the Airspace Technology Demonstration-2 (ATD-2) Integrated Arrival / Departure / Surface (IADS) activity also introduces ATC/Ramp Control DSTs to support more efficient surface operations [4]. For example, a DST will provide Ramp

Controllers with pushback advisories designed to reduce long departure queues [4] and allow aircraft to remain at the gate, with engines off, saving fuel and reducing emissions. In both of these examples (i.e., STBO and ATD-2 IADS), ATC and/or Ramp Control will use DSTs to manage aircraft to meet the schedule requirements. In the near-term timeframe, the flight deck is not expected to be equipped with avionics displays that support schedule conformance.

B. *Far-Term STBO Operations*

Leveraging NextGen STBO operations, farther-term concepts have been envisioned that require pilots to meet a required time of arrival (RTA) at traffic-flow constraint points on the airport surface [2] [10]. In order to enable aircraft to meet RTAs, more coordination between ATC/Ramp Control and the flight deck will be required, as well as flight deck avionics to facilitate schedule conformance.

These farther-term STBO concepts can be considered along a continuum of increasing temporal certainty as the number of traffic-flow constraint points along the taxi route increase. For example, at one end of the continuum, operations in which only a spot-release time is scheduled offer the least amount of temporal certainty about when an aircraft will arrive at the departure queue. To increase temporal certainty, aircraft will be expected to meet RTAs at *intermediate* traffic-flow constraint points between the spot and the departure queue as well (e.g., taxiway merge points, active-runway crossings) [11]. At the other end of the continuum are four-dimensional trajectories (4DTs) which define an expected location (*x,y coordinates or latitude, longitude*) at all times, *t*, along the taxi route (with *altitude*, being fixed) [2] [10]. These 4DT operations afford the highest temporal certainty at *all* points along the taxi route, and at the departure runway. This allows for more efficient crossing of traffic at taxiway intersections as well as efficient departure runway queues.

C. *Far-Term 4DT STBO Operations*

In the far-term, 4DT STBO concept, each aircraft is assigned a conflict-free, four-dimensional trajectory (4DT). Fig. 1 depicts 4DT taxi operations where each aircraft on the surface has been assigned a conflict-free 4DT route which has an expected position (black dot) at all times along the taxi route. Each 4DT also has an error tolerance (i.e., allowable deviation) to which the aircraft must conform to maintain compliance to the 4DT and avoid conflict with other traffic. In order to aid pilots in safely complying with the increased-timing requirements of a 4DT taxi clearance, advanced flight deck equipage is required.

In the far-term 4DT STBO environment, ATC is expected to use surface management tools/automation to generate conflict-free 4DT taxi clearances and to monitor conformance. One example of such a tool is the Taxi Routing for Aircraft: Creation and Controlling (TRACC) system developed by the German Aerospace Center (DLR). TRACC is a research prototype surface management system that generates conflict-free 4DT speed profiles [12].

Two concepts of operations (ConOps) for 4DT surface operations have been developed. The National Aeronautics and Space Administration's (NASA) far-term STBO ConOps [2] describes the far-term vision for surface operations in the U.S. National Airspace System. A harmonized ConOps for 4DT taxi operations, which considers the different surface-management practices and policies that currently exist in the United States and Europe, was developed jointly by NASA and DLR [10]. The harmonized 4DT ConOps represents a future surface operations concept that could be implemented in the United States and Europe.

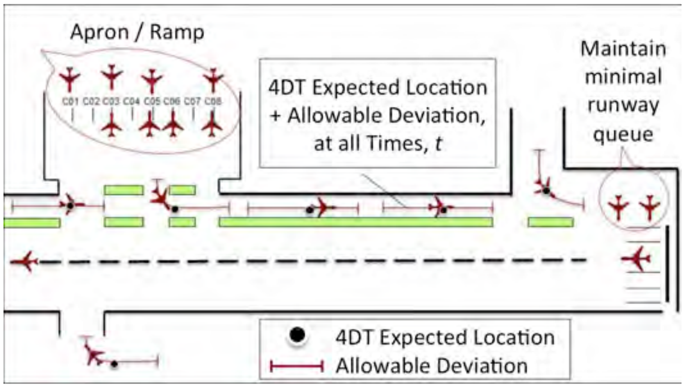

Fig. 1. 4DT STBO with expected position (black dot) and tolerance band (red band) for each aircraft [10].

This paper will describe a pilot-in-the-loop simulation in which far-term, 4DT surface operations were explored from the perspective of the flight deck. First, previous flight deck simulations that describe the need for a human-centered flight deck display algorithm are reviewed followed by a previous 4DT proof-of-concept flight deck simulation study that demonstrated the feasibility of the 4DT concept. Finally, the results of the present study, in which 4DT display formats were compared, are reported. The present study considered more robust operating conditions than the first 4DT study, including 4DT speed profile updates during taxi and a range of realistic taxi speeds.

III. FLIGHT DECK DISPLAYS FOR STBO

A. *Speed-Based Taxi Clearances on the Flight Deck*

Pilot-in-the-loop flight deck simulations have been conducted to assess pilots' ability to meet the time requirements of the far-term STBO environment (e.g., RTAs at traffic-flow constraint points) [11]. During taxi, the only control mechanism that the pilot has to reach a certain location on the airport surface at a certain time is through the control of aircraft speed. Thus, *speed*-based taxi clearances and their effect on pilots' ability to meet an RTA were explored in two studies (results summarized in Table I).

First, pilots were asked to follow a speed command, issued as part of the taxi clearance, and displayed on the PFD, to meet the RTA at the departure queue (Expt. 2 in [11]). RTA error, that is, the difference between the required time of arrival and the aircraft's actual time of arrival, was unacceptably large and considered not precise enough for STBO operations.

In a second study (Expt. 3 in [11]), pilots were instructed to taxi within +/-1.5 kts of the verbally commanded speed and accelerate/decelerate at 2 kts/sec (a verbal "check speed" alert was delivered when the ground speed exceeded the +/-1.5 kt range for more than a continuous 5-sec period). While RTA error was smaller as a result of imposing a speed bound requirement (i.e., +/-1.5 kts), pilots spent an excessive amount of time head down (i.e., "eyes-in") tracking the ownship's speed readout. As a result, 14 out of 18 pilots rated the procedure as 'unsafe'.

The findings of these two studies suggest that providing pilots with a commanded taxi speed, as a means for meeting an RTA along the taxi route, is not sufficient for both safety and performance. There is a need for a flight deck display/tool to aid pilots in safely and precisely meeting the expected precision timing requirements of the STBO concept.

B. *Human-Centered Flight Deck Display Algorithm*

In a third study, the flight deck was equipped with an error-nulling speed algorithm which displayed the current advised-speed, calculated by dividing the remaining distance by the remaining time (Expt. 4 in [11]). The speed algorithm compensated for speed-maintenance deviations by dynamically updating the current advised speed needed to meet the RTA (e.g., if the aircraft's speed slowed, then the advised speed gradually increased to compensate for that speed deviation, and vice versa). Results showed that the error-nulling speed algorithm supported both safety and performance. RTA error, the variability in arrival time at the RTA point, was reduced for greater timing precision. In addition to the increase in timing precision, the error-nulling algorithm did not compromise out-the-window attention because pilots were not required to track the speed precisely. These three flight deck STBO studies are summarized in Table I.

TABLE I. SUMMARY OF FLIGHT DECK STBO RESEARCH

Flight Deck Displays for STBO (Experiments 2, 3, and 4 in [11])			
Procedural Instructions	***Flight Deck Equipage***	***Required Time of Arrival (RTA) Performance (Error)***	***Safety (Head-Down Time)***
Taxi at Commanded Speed $^{\text{Expt. 2}}$	No additional equipage.	RTA error unacceptably large and not precise enough for STBO operations.	Slight increase in head-down time.
Taxi within +/-1.5 kts of Commanded Speed $^{\text{Expt. 3}}$	No additional equipage.	Small RTA error; Good (precise) RTA performance.	Excessive head-down time; rated "unsafe".
Taxi at Advised Speed $^{\text{Expt. 4}}$	Error-Nulling Speed Algorithm	Small RTA error; Good (precise) RTA performance.	Head-down time rated "acceptable".

Note: Red shading: Unacceptable RTA performance or unsafe eye-tracking results; Green shading: Good RTA performance or safe eye-tracking results.

C. *4DT Flight Deck Displays*

In 4DT surface operations, each aircraft will be assigned a conflict-free 4DT taxi clearance and will have an expected *x,y* location at *all* times *t*, along the taxi route. While the error-nulling algorithm display, described in the previous section, safely and precisely supported RTA conformance at a specific point (e.g., the runway queue), there was considerable variability in aircraft speed along the route which led to a lack of predictability about aircraft location along the route. To support pilots in conforming to an expected 4DT location at *all* times along the route, a 4DT flight deck display is required. Work has been done to develop flight deck Airport Moving Map (AMM) displays (presented on flight deck avionics or on an Electronic Flight Bag, EFB) that aid pilots in safely conforming to the time/speed requirements of 4DT taxi operations, where there is an expected position at *all* times along the route, and to support traffic conflict detection and avoidance in the 4DT operations [13] [14] [15].

A previous study explored pilots' ability to safely conform to a 4DT, while taxiing under manual control, with the aid of a 4DT flight deck display [15]. In two 4DT taxi conditions, the AMM, which also displayed the 4DT speed, was augmented to graphically show: 1) 4DT reference markers that represented the expected ownship position, and 2) the allowable tolerance (i.e., required conformance) around that expected position. The reference markers and tolerance band moved, dynamically, along the taxi route according to the 4DT speed profile. This 4DT flight deck display was designed to be a status-at-a-glance display to maximize out-the-window time and enable strategic use so pilots are not compelled to track speed.

Two time-based tolerance-band sizes were compared, +/-15 sec and +/-30 sec. Tolerance-band sizes were converted to distance, based on the assigned speed (i.e., at 15 kts, +/-380 ft and +/-760 ft, respectively). 4DT taxi movement started at the ramp departure spot (Airport Movement Area (AMA) entrance) and ended at the entrance to the departure queue. The assigned 4DT speed (i.e., 14, 15, or 16 kts) was held constant throughout each trial.

Results showed that both 4DT tolerance bands afforded more than 99% conformance to the 4DT allowable tolerance across the taxi route and reduced variability in time of arrival at the departure queue. An analysis of eye-tracking data showed that pilots spent less time scanning out-the-window in both the +/-15 sec and +/-30 sec allowable-tolerance conditions, 62.1% and 65.6%, respectively, as compared to the current-day operations condition (81.3%) where the taxi clearance did not include time or speed requirements. However, pilots rated the head-down time as 'acceptable' in all conditions. The 4DT displays afforded high conformance with an increase in predictability throughout the entire taxi route and at the departure queue.

IV. PRESENT STUDY

The purpose of the present study was to compare the effect of 4DT display formats on conformance, safety, and pilot workload ratings, while taxiing with a 4DT clearance, under manual control. In this far-term STBO concept, 4DT taxi clearances were generated as a single clearance, from gate to runway queue, designed to reduce ramp-area congestion and eliminate the bottleneck associated with transitioning from the ramp to the AMA. The display formats examined included:

1) Defined-Tolerance Display Format: In the *defined*-tolerance display format, a graphical representation of the expected 4DT ownship position, with a distance-based allowable-tolerance band, was depicted on the AMM. Pilots were instructed that they were in compliance with the 4DT clearance when their aircraft was within the tolerance band. Two tolerance-band sizes were tested +/-164 ft and +/-405 ft.

2) Undefined-Tolerance Display Format: In the *undefined*-tolerance display format, the expected ownship position was displayed graphically on the AMM, with no indicated allowable-tolerance bounds. A 4DT indicator "dot", moving along the taxi route on the AMM, displayed the expected 4DT location. Pilots were instructed to use their best judgment to decide how closely to track the 4DT indicator during taxi and were free to taxi either ahead of or behind the 4DT indicator.

The present study also introduced a wider range of real-world conditions than the previous 4DT flight deck study [15] in which the 4DT speed remained constant for duration of the route. In the present study, the 4DT speed was updated during taxi (two or five times, per trial), and a wider range of speeds were used to create 'slow' and 'fast' taxi speed scenarios.

V. METHOD

A. Participants

Twelve commercial airline pilots (11 Captains, 1 First Officer) participated in the simulation. Eleven pilots were current, and one Captain was recently retired. The mean pilot age was 56 years (range: 50 to 65 years). All twelve pilots had taxi experience; mean flight hours logged as Captain in command (n=11) was 9,730 hours and the First Officer reported logging more than 5,000 hours. Pilots' had varying degrees of familiarity with the airport layout. Each pilot was paired with a First Officer, who was a member of the research team, to form a two-pilot crew. This same First Officer was paired with all 12 participants and provided navigation and traffic awareness support in a consistent manner to each pilot participant.

B. Airport and Terminal Area Simulator (ATAS)

The study was conducted in the Airport and Terminal Area Simulator (ATAS) in the Human-Centered Systems Laboratory (HCSL) at NASA Ames Research Center. The ATAS is a modified B737-NG cockpit equipped with an unobtrusive four-camera eye-tracking system (Smart Eye Pro; Smart Eye AB, Goteborg) to measure pilot gaze location.

The airport environment was the Charlotte Douglas International Airport (KCLT) with high-visibility and distant fog/haze conditions. The forward, out-the-window scene was depicted on four LCD displays, with a total horizontal viewing angle of 140 deg. The physical and taxi handling characteristics of the aircraft were that of a mid-size, narrow-body aircraft. The flight deck was equipped with a PFD (inactive for this study), an engine-indicating and crew-alerting system (EICAS), and an airport moving map (AMM) in place of the traditional navigation display (ND).

C. 4DT Taxi Displays

The Airport Moving Map (AMM) depicted the airport layout, in perspective and in track/heading-up format, to aid pilots in airport navigation. As depicted on the AMM, the ownship aircraft was shown as a white chevron icon and scaled to 99 ft long by 99 ft wide. Airport traffic within 1,250 ft of the ownship were updated in real time. For analysis purposes, the precise x,y location of the ownship aircraft was a point near the middle of the ownship icon (specifically, 45 ft back from the front apex of the aircraft icon/chevron and centered laterally). The cleared-to-taxi route, from the terminal to the departure runway, was displayed in dark magenta (Fig. 2).

In the two defined-tolerance display conditions, the expected 4DT ownship position was represented by two horizontal reference markers (shown in front of the ownship in Fig. 2). The light pink segment that overlays the magenta route represents the allowable tolerance from the expected ownship position. Two allowable-tolerance sizes were tested in this study, +/-164 ft and +/-405 ft (+/-405 ft depicted in Fig. 2).

In the undefined-tolerance display condition, the 4DT indicator that indicated the expected ownship position was depicted as a light pink dot (Fig. 3), with no indicated allowable-tolerance bounds. The diameter of the 4DT dot icon as scaled on the AMM was 72 ft.

Current ground speed and heading ("GS 15" kts and 087 deg in Fig. 2) were displayed at the top of the map. The cleared-to-taxi route (RWY 18L via Mike, Charlie in Fig. 2), 4DT start time (12:05:30), current 4DT speed (15 kts), and Target Takeoff Time (TTOT) (12:11:00) were shown textually in magenta, below the map.

Taxi and pushback clearances were delivered to the flight deck via DataComm instead of by voice. The DataComm touchscreen interface was located aft of the throttles between the two pilots. Pilots received the DataComm taxi clearance while at the gate, prior to receiving the DataComm pushback clearance. All DataComm messages were accompanied by an auditory chime and were read aloud by the First Officer.

D. Experimental Design

Three 4DT display conditions were tested: two distance-based defined-tolerance bands, +/-164 ft and +/-405 ft, and one undefined-tolerance condition where the expected ownship position, according to the 4DT speed profile, was represented graphically as a dot on the AMM.

As mentioned previously, the DLR TRACC system is an example of a prototype surface management system that generates 4DT speed profiles from the gate to the runway, monitors 4DT conformance, and issues speed/routing updates to resolve traffic conflicts [12]. Several parameters used in the present study were based on those used in the TRACC system.

The length of the smaller of the two defined-tolerance conformance bounds was +/-164 ft, which matches a parameter used by TRACC algorithms (164 ft = 50 m). TRACC algorithms have proposed to use +/-164 ft as a distance threshold for conformance monitoring. That is, the TRACC system monitors each aircraft's distance from its expected 4DT location during taxi and, if that distance exceeds +/-164 ft, adapts the assigned speed and assesses schedule and/or traffic conflicts. If determined to be necessary, the system may issue speed updates or re-routing guidance to resolve any schedule or traffic conflicts. While the +/-164 ft bound is used as an internal conformance parameter to the TRACC system, the present pilot-in-the-loop simulation implemented +/-164 ft as a 4DT conformance requirement for pilots, presented graphically on the flight deck AMM, while they taxied the aircraft under manual control.

The length of the larger of the two defined-tolerance conformance bounds used in this study was +/-405 ft, which approximates the size of the smaller time-based tolerance band (at 16 kts) used in the previous 4DT study (i.e., +/-15 sec at 16 kts = +/-405 ft) [15]. In the previous study, the +/-15 sec tolerance band afforded 99% compliance with the 4DT.

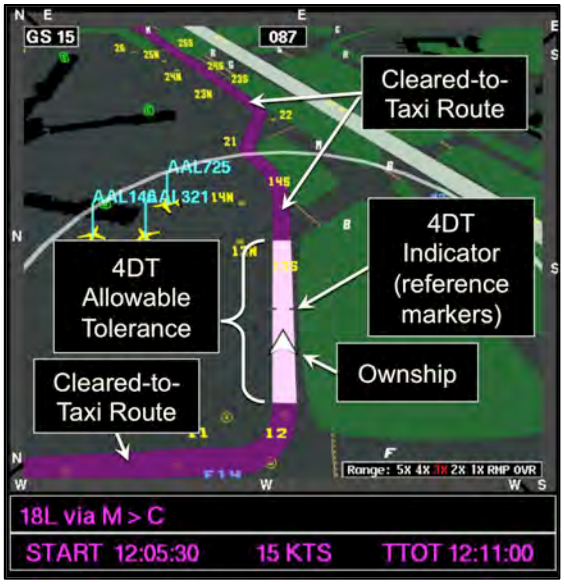

Fig. 2. Airport Moving Map (AMM) in the *defined*-tolerance display format with 4DT indicator (black reference lines in front of the ownship) and +/-405 ft allowable tolerance (light pink band).

Fig. 3. Airport Moving Map (AMM) in the *undefined*-tolerance display format with 4DT indicator (light pink dot in front of the ownship).

In the undefined-tolerance condition, the expected 4DT location was represented graphically on the AMM as a dot, with no indicated allowable-tolerance bounds (see Fig. 3). The purpose of this condition was to assess pilots' conformance to a 4DT when an allowable tolerance was not depicted graphically on the AMM or defined through procedures.

The flight deck received 4DT speed updates at predetermined locations in each trial (two or five updates, per trial). The number of 4DT speed changes was chosen based on the frequency of updates observed in a TRACC simulation. Pilots were alerted to 4DT speed changes by an auditory tone; concurrently, the textual 4DT speed display below the AMM (e.g., 8 kts in Fig. 3) also updated. On average, speed changes were located 2,211 ft, or 101 sec, apart in the two speed-change trials, and 1,105 ft, or 51 sec, apart when the trial included five 4DT speed changes. Two speed changes yielded three taxi segments (two in the Ramp and one in the AMA), while five speed changes yielded six taxi segments (three in the Ramp and three in the AMA).

Taxi speeds ranged between 8 kts and 25 kts in this study. Taxi speeds were assigned to taxi segments to create 'slow' and 'fast' 4DT speed conditions. In the 'slow' speed condition, the average 4DT taxi speed was 10 kts in the Ramp area and 16 kts in the AMA; and in the 'fast' speed condition, 13 kts and 22 kts in the Ramp and AMA areas, respectively. The 4DT profile speeds, assigned to each taxi segment, are shown in Table II. These 4DT speeds represent the speed at which the 4DT indicator (i.e., tolerance band or dot) moved along the taxi route in each segment. Pilots were not required to track speeds precisely, but rather, maintain conformance to the 4DT. In the defined-tolerance condition, this meant keeping the ownship aircraft icon within the tolerance band. In the undefined-tolerance condition, pilots defined conformance as they saw fit. The three display conditions were factorially crossed with 4DT speeds and 4DT speed changes to create a 3 (4DT display condition) by 2 ('slow' or 'fast' 4DT speed) by 2 (two or five 4DT speed changes) within-subjects design, with a total of 12 experimental trials.

TABLE II. 4DT PROFILE SPEEDS BY TAXI SEGMENT (IN KTS)

Avg. 4DT Speed	**4DT Speed Changes**	**Ramp**				**Airport Movement Area (AMA)**			
		Taxi Segment				***Taxi Segment***			
		1	***2***	***3***	***Avg.***	***1***	***2***	***3***	***Avg.***
Slow	2	11	9	-	**10**	16	-	-	**16**
Slow	5	11	8	11	**10**	14	19	15	**16**
Fast	2	11	15	-	**13**	22	-	-	**22**
Fast	5	11	15	13	**13**	20	25	21	**22**

[a] These four trials were randomized and repeated in each of the three 4DT display-format conditions.

The 4DT taxi route started in the ramp area, near the terminal, as described in the NASA/DLR harmonized 4DT ConOps [10], and ended near the queue area at the departure runway. The 4DT taxi clearance was a continuous clearance from the terminal, through the ramp, and to the runway. Two taxi routes at KCLT airport were used in the study: Terminal D to runway 18C and Terminal A to runway 18L. The average distance/duration of the two routes was 6,633 ft / 306 sec. The ramp accounted for approximately 57% of the distance in each taxi route. The two routes were matched in terms of the number of turns. In accordance with airport surface operations standard operating procedures (SOPs), the 4DT indicator did not slow or speed-up for turns in the ramp area, but did slow to 14 kts for turns in the AMA. 4DT speed profiles used a 1 kt/sec acceleration/deceleration rate when accelerating from 0 kts at the beginning of the route and for each speed change.

The three 4DT display conditions were blocked and their order was assigned using a Latin square. The order of the four trials within each display condition (slow speed/two speed changes; slow speed/five speed changes; fast speed/two speed changes; and fast speed/five speed changes) were randomized.

E. *Procedure*

Before beginning the experimental trials, pilots taxied two general simulator-familiarization trials, and a third 4DT-familiarization trial, in which they were introduced to the 4DT taxi concept. In this third trial, pilots taxied with the undefined-tolerance 4DT indicator (dot). Prior to each of the three experimental blocks, pilots received training specific to that particular 4DT display format and taxied a practice trial for that condition (not included in the analysis).

In all trials, pilots were instructed that their first priority was to maintain safety, never taxi faster than they would in an actual aircraft full of passengers, and remain eyes-out and taxi with the same regard for passenger safety and comfort as in actual operations. In the two defined-tolerance conditions (+/-164 ft and +/-405 ft), pilots were told that their first priority was to maintain safe, eyes-out taxi, and their second priority was to comply with the 4DT as best they could. They were instructed that they would be in compliance with the 4DT if their aircraft was within the allowable deviation (i.e., the light-pink band displayed on the AMM). Pilots were not required to track the center reference markers, or the 4DT speed, precisely. If they found themselves outside of the tolerance band, they were instructed to recapture it as quickly, but as safely, as possible (i.e., with appropriate safe, typical acceleration or deceleration).

In the undefined-tolerance condition, the expected ownship position was represented by a 4DT indicator (dot). Pilots were told that while taxiing close to the 4DT indicator (dot) is expected to increase overall airport efficiency and minimize delay for them, and other aircraft, they were not expected to track the 4DT indicator (dot) precisely. Pilots were instructed that they were to decide how "close is close enough" to taxi to the dot, and that they could taxi ahead of, or behind, the 4DT dot. No procedural instructions about allowable-tolerance bounds were provided in this condition.

Each trial started at the gate, where the flight deck received the taxi clearance via DataComm. The taxi clearance included a continuous clearance from the terminal to the assigned departure runway, as well as, schedule information: Target Off-Block Time (TOBT; expected pushback-clearance time), forward taxi time (the time at which taxi begins; where the 4DT indicator begins to accelerate and move along the taxi route, defined by the 4DT speed profile), and the Target Takeoff Time (TTOT; estimated takeoff-clearance time, provided as an advisory only).

The taxi routing and schedule information were automatically loaded into the flight deck avionics and appeared in cyan, until forward taxi time. After completing pushback, the AMM changed from overview mode to track-up perspective and pilots positioned their ownship icon within the bounds of the tolerance band, or near the 4DT indicator (dot) in the undefined-tolerance condition. At forward taxi time (beginning of taxi) a tone sounded, the 4DT indicator (tolerance band or dot) began to accelerate, and clearance information on the AMM turned from cyan to magenta.

The 4DT taxi route ended at the entrance to the queue area near the departure runway. As the ownship crossed that queue-entry point, the 4DT tolerance band or dot disappeared and speed guidance was blanked. Pilots were instructed to continue taxiing safely into the queue area, follow the aircraft in front of them, and follow ATC instructions.

VI. RESULTS

This study compared the effect of 4DT display formats on distance from 4DT indicator, conformance to defined-tolerance band, eyes-out time, and pilot ratings of safety, and workload. The defined-tolerance display format was tested at two levels +/-164 ft and +/-405 ft. For two subjects, one trial each was removed from the data in the +/-164 ft tolerance-band condition because those pilots unexpectedly responded to a piece of traffic crossing their taxi route by bringing their aircraft to a stop; those two data points were replaced with the mean value for that condition so as to not affect analysis.

A. Distance (Absolute Value) from 4DT Indicator

Across the three 4DT display conditions (+/-164 ft, +/-405 ft, and dot), the distance between the expected 4DT location, that is, the middle of each tolerance band, or the middle of the dot, and the ownship aircraft was recorded during taxi. The precise location of the ownship aircraft was determined by a point near the middle of the ownship icon.

As would be expected, the absolute value of the distance was shorter in the smaller defined-tolerance condition than in the larger tolerance-band condition. Pilots taxied an average of 67.51 ft from the center of the +/-164 ft tolerance band, and 92.22 ft from the center of the +/-405 ft tolerance band. Distance from the 4DT indicator was highest in the undefined-tolerance condition, where pilots taxied an average of 94.72 ft from the center of the 4DT dot.

Distance between the ownship and the expected 4DT location was analyzed in a 3 (4DT display) by 2 (4DT speed) by 2 (number of 4DT speed changes) repeated-measures ANOVA, which revealed a 4DT speed by number of 4DT speed changes interaction, $F(1,11)=5.13$, $p<.05$. As shown in Fig. 4 (bottom), in the five-speed change condition, distance between the ownship and expected 4DT location was greater when the average 4DT speed was 'slow' ($M=99.73$ ft, $S.E.=18.09$ ft) than when 4DT speed was 'fast' ($M=72.95$ ft, $S.E.=7.18$ ft), $p<.05$. However, the taxi strategy used in the undefined conformance condition, contributed to this interaction. At least one pilot, who experienced the undefined conformance condition before either of the tolerance-band conditions, taxied ahead of the 4DT indicator. The pilot used this strategy to ensure that any delay the ownship aircraft might encounter would not prevent the pilot from reaching the end of the 4DT taxi route at the departure queue at precisely the same time as the 4DT indicator. The pilot maintained a position well in front of the 4DT indicator in the 'slow' speed, five speed-change trial.

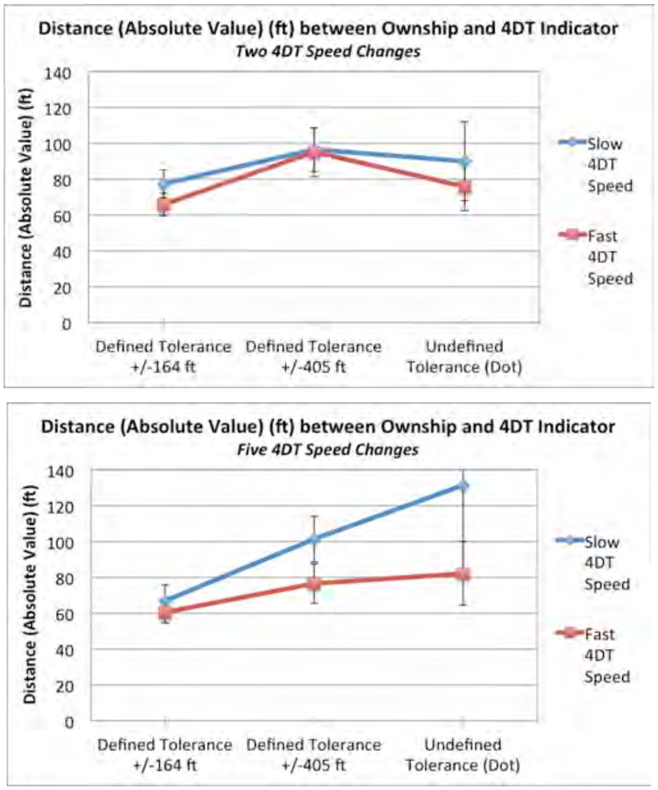

Fig. 4. Distance (absolute value, ft) between the ownship and the expected 4DT location in the two (top) and five (bottom) 4DT speed change conditions.

Another way to assess distance from the 4DT indicator is shown in Fig. 5, which shows the percentage of time the ownship aircraft was within a given distance from the expected 4DT location, for each of the three 4DT display conditions. Fig. 5 shows that pilots spent more time taxiing closer to the expected 4DT location in the smaller (+/-164 ft) tolerance-band condition than in either the larger (+/-405 ft) tolerance band or undefined-tolerance conditions.

It should be noted that because the order of the three 4DT display conditions was counterbalanced, two-thirds of pilots in the study (n=8) were exposed to one, or both, of the defined-tolerance conditions prior to taxiing with the dot. Exposure to taxing with defined-conformance bounds may have created an expectation about the 'acceptable' distance for pilots to taxi from the 4DT dot.

Predictability of the aircraft's location can be considered by examining the distance from the 4DT indicator for a given confidence interval, for example, 95%. As can be seen in Fig. 5, 95% of the time, pilots maintained their aircraft within +/-175 ft when taxiing with the smaller defined-tolerance band, +/-250 ft with the larger defined-tolerance band, and +/-300 ft in the undefined-tolerance (dot) condition.

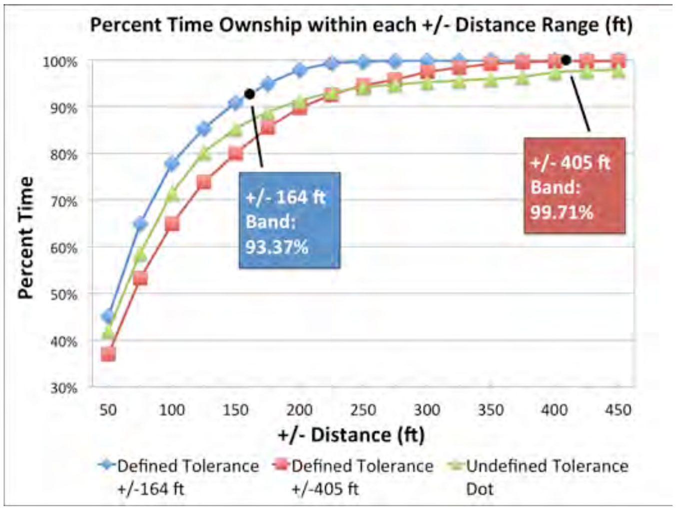

Fig. 5. Percentage of time the ownship taxied within each +/- distance range from the expected 4DT location. Circles denote conformance percent for the two defined-tolerance display formats (+/-164 ft and +/-405 ft).

B. *Conformance to Defined-Tolerance Band*

Conformance to the defined-tolerance band is measured as the percentage of time that the aircraft icon was located within the displayed tolerance band, and therefore can only be examined in the two defined-tolerance display format conditions. The percentage of time the ownship aircraft was in conformance with each of the defined-tolerance bands (+/-164 ft and +/-405 ft) was examined. The larger (+/-405 ft) tolerance band afforded very high conformance, 99.71%, across the entire taxi route. The smaller (+/-164 ft) tolerance band, however, resulted in lower conformance, 93.37%. A 2 (4DT defined-tolerance band) by 2 (4DT speed) by 2 (number of 4DT speed changes) repeated-measures ANOVA indicated that only the 4DT tolerance band main effect (i.e., +/-164 ft vs. +/-405 ft) was significant, $F(1,11)=7.95$, $p<.05$.

It should be noted that some pilots were observed using a specific strategy to maintain conformance to the +/-164 ft band. These pilots positioned their ownship on the front edge of the +/-164 ft tolerance band, so that just the back edge of the ownship icon overlapped with the band, as a strategy to guard against falling behind when the 4DT speed increased while still maintaining the required 4DT conformance. Because the location of the ownship was determined by a point near the middle of the ownship icon, distance from the middle of band to the ownship's precise location would have been greater than +164 ft when pilots adopted this strategy. If the conformance measurement is adjusted on the front end of the band to allow for *any* part of the ownship icon overlapping the band, 4DT conformance increases from 93.37% to 97.35%. (Note: This value represents an upper-bound estimate, since it would "score" the aircraft as being in conformance even if a single pixel of the ownship icon was on the conformance band.)

Although conformance bounds were not specified in the undefined-tolerance condition, the percentage of time the ownship taxied within +/-164 ft and +/-405 ft of the expected 4DT location was examined for comparison. The ownship aircraft was within +/-164 ft of the expected 4DT location (i.e., center of the dot) 87.48% of the time, and within +/-405 ft of the expected 4DT location 97.47% of the time. For the four pilots who experienced the undefined-tolerance display format first, the ownship aircraft was within +/-164 ft of the expected 4DT location (i.e., the center of the dot) 77.80% of the time, and within +/-405 ft of the 4DT location 92.95% of the time.

C. *Eyes-Out Time*

The effect of 4DT display format on safety was assessed by measuring the time spent looking out-the-window during taxi. Fig. 6 shows the percentage of the time spent scanning out-the-window (eyes-out) and the flight deck instrument panel (eyes-in) between the ramp and the departure queue. Pilots who had undefined data (e.g., looking at something other than a defined areas of interest such as throttles or tiller) or unreliable data (e.g., eyes closed, blinks) for 20% or more of a trial were excluded from these analyses, yielding a subset of nine subjects.

Pilots spent more time scanning out-the-window in the larger +/-405 ft defined-tolerance condition (*M*=65.53%, *S.E.*=2.23%) and undefined-tolerance (dot) condition (*M*=65.23%, *S.E.*=2.16%), than in the smaller +/-164 ft tolerance-band condition (*M*=61.84%, *S.E.*=2.23%); however, the main effect of 4DT display was not significant, *p*=.069.

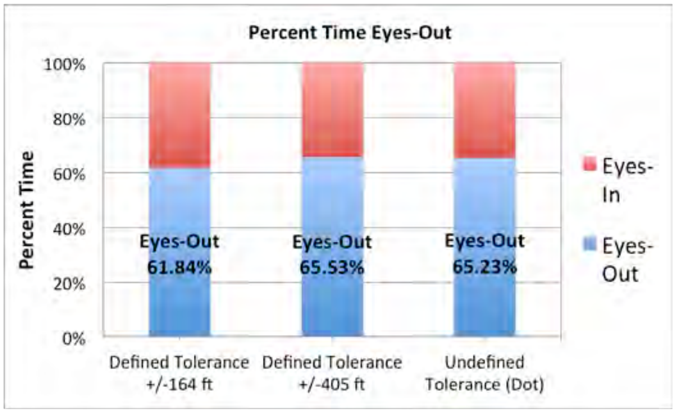

Fig. 6. Percentage time spent scanning out-the-window (eyes-out time) vs. flight deck displays (eyes-in time) in each 4DT display condition.

A 3 (4DT display) by 2 (4DT speed) by 2 (number of 4DT speed changes) repeated-measures ANOVA showed that pilots spent more time scanning out-the-window in the two speed-change condition (*M*=65.31%, *S.E.*=1.73%) than in the five speed-change condition (*M*=63.09%, *S.E.*=1.87%), *F*(1,8)=5.24, *p*=.051. When pilots received a 4DT speed update during taxi, the speed display on the bottom of the AMM was updated. More frequent speed changes increased the amount of eyes-in time.

D. *Subjective Ratings Eyes-In Time*

To understand pilots' experience with the eyes-in time required to taxi with 4DT displays, pilots' assessment of the eyes-in time was explored in two questions. First, following each trial, pilots were asked, "During this trial, how often did you find yourself focusing on the speed and/or time displays when you should have been paying attention to the external taxiway environment?", on a 5-point scale, where 1=Rarely, 2=Seldom, 3=Sometimes, 4=Frequently, and 5=Most of the Time. Pilot-ratings showed that they more often focused on the flight deck displays when they taxied with the smaller tolerance band (+/-164 ft) than with the larger tolerance band (+/-405 ft) or the undefined-tolerance (dot) display.

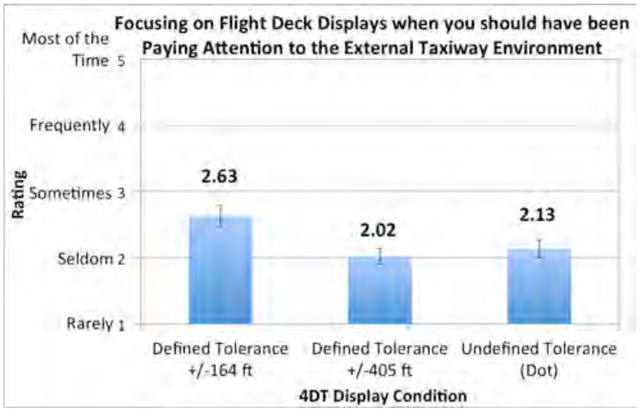

Fig. 7. Post-trial ratings of eyes-in time in each 4DT display condition.

Pilots were also asked to rate the acceptability of the eyes-in time required for each 4DT display format using a 5-point scale where 1=Very Unacceptable, 2=Unacceptable, 3=Borderline, 4=Acceptable, and 5=Very Acceptable. A repeated-measures ANOVA revealed an effect of 4DT display condition, *F*(2,22)=4.53, *p*<.05.

As shown in Fig. 8, pilots rated eyes-in time as more acceptable when taxiing with the +/-405 ft tolerance band (*M*=4.25, *S.E.*=0.18) than with the +/-164 ft band (*M*=3.33, *S.E.*=0.38), *p*<.05, or the undefined-tolerance format (*M*=3.75, *S.E.*=0.18), *p*=.053. The difference between the +/-164 ft band and undefined-tolerance format was not significant.

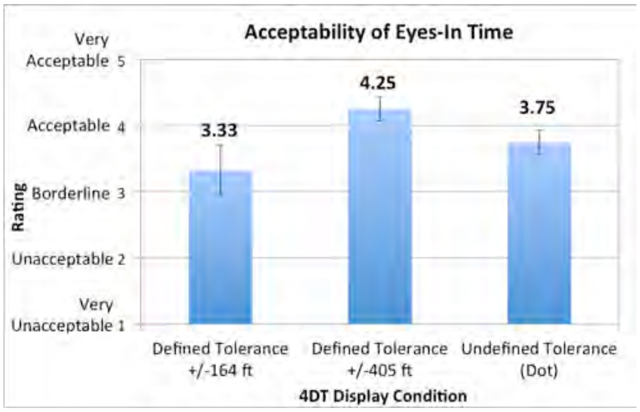

Fig. 8. Acceptability of eyes-in time in each 4DT display condition.

E. *Subjective Ratings of Safety*

On a post-study questionnaire, pilots rated the "safety of taxiing with each of the 4DT display formats" using a 5-point scale where 1=Very Unsafe, 2=Somewhat Unsafe, 3=Borderline, 4=Somewhat Safe, and 5=Very Safe.

A repeated-measures ANOVA showed an effect of 4DT display condition, *F*(2,22)=3.54, *p*<.05. As shown in Fig. 9, pilots' perceived safety was higher when taxiing with the +/-405 ft tolerance band (*M*=4.42, *S.E.*=0.23) than the +/-164 ft tolerance band (*M*=3.58, *S.E.*=0.36), *p*<.05, and the undefined-tolerance (dot) format (*M*=3.92, *S.E.*=0.19), *p*=.053. The difference between the +/-164 ft tolerance band and the undefined-tolerance (dot) format was not significant.

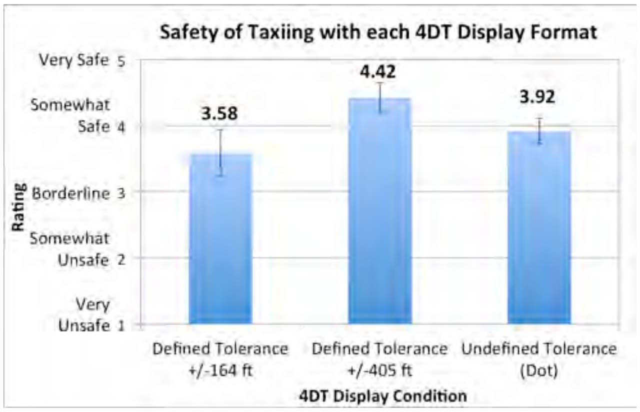

Fig. 9. Post-study safety ratings of each 4DT display condition.

F. *Subjective Ratings of Workload*

On a post-trial questionnaire, pilots rated the "overall workload required to successfully taxi" each trial using a 5-point scale where 1=Low, 3=Neutral, and 5=High.

As shown in Fig. 10, pilots perceived workload to be higher when taxiing with the +/-164 ft tolerance band (*M*=2.77, *S.E.*=0.15), than the +/-405 ft tolerance band (*M*=2.50, *S.E.*=0.13), or the undefined-tolerance (dot) format (*M*=2.54, *S.E.*=0.13). However, a 3 (4DT display) by 2 (4DT speed) by 2 (number of 4DT speed changes) repeated-measures ANOVA revealed no significant differences.

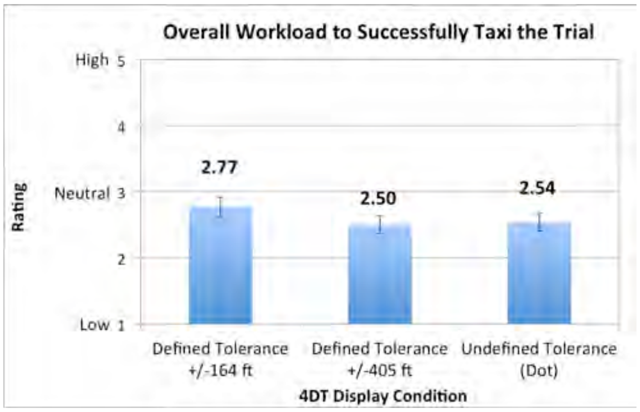

Fig. 10. Post-trial ratings of overall workload in each 4DT display condition.

VII. CONCLUSIONS

The present study compared the effect of 4DT display formats on distance from 4DT indicator, conformance to the defined-tolerance band, eyes-out time, and pilot ratings of safety and workload, while taxiing with a 4DT clearance under manual control. In the defined-tolerance display format, a graphical representation of the expected 4DT ownship position, with a distance-based allowable-tolerance band, was depicted on the AMM. Two tolerance-band sizes were tested: +/-164 ft and +/-405 ft. In the undefined-tolerance display format, the expected ownship position was displayed graphically on the AMM, with no indicated tolerance bounds. The 4DT speed was updated during taxi (two or five times, per trial), and 4DT speeds assigned to create 'slow' and 'fast' taxi speed scenarios.

Conformance to the defined-tolerance bands used in this study can be thought of as the aircraft's compliance to the "defined airport real estate" of the assigned 4DT clearance, which is used by the ATC surface management system to manage taxiing aircraft. Any deviation from that assigned conformance band could, but may not necessarily, trigger a new 4DT clearance assignment. If the aircraft is "out of conformance" (i.e., not on the 4DT defined-tolerance band), the ATC surface management system would use the aircraft's current location to recalibrate or issue a new 4DT clearance.

Conformance to a 4DT taxi clearance and the predictability of an aircraft's location relative to a specified 4DT location are related metrics, but not identical. This study showed that the larger +/-405 ft defined-tolerance band yielded higher conformance levels than the smaller +/-164 ft band, with pilots keeping their aircraft within the specfied and displayed conformance bounds to a greater degree in the larger (99.71%) than the smaller defined-tolerance band (93.37%). However, in terms of being able to predict the location of the aircraft compared to the expected 4DT location, the smaller defined-tolerance band resulted in pilots keeping their aircraft closer to the 4DT location (for both average distance, and for a given confidence interval, e.g., 95%). More research is required to understand the system-level trade-offs and interactions associated with aircraft conformance requirements and location predictability with 4DT clearances.

Although it is not yet known if conformance to a tolerance-band size of +/-405 ft will be precise enough to support the requirements of 4DT surface operations, the larger tolerance band afforded two positive findings. The larger tolerance band yielded more "eyes out-the-window" time than the smaller tolerance band, and the "eyes-in" time associated with the larger tolerance band was rated as more acceptable than the smaller tolerance band. Pilots also rated taxiing with the larger 4DT tolerance band as safer than the smaller 4DT band.

On average, pilots taxied at a similar distance from the expected 4DT location with the undefined-tolerance (dot) display as they did with the larger +/-405 ft tolerance band display. However, the range of distances that pilots taxied from the expected 4DT location was larger in the undefined-tolerance (dot) condition because it allowed pilots to interpret 'conformance' idiosyncratically and employ different taxi strategies (e.g., some pilots may taxi a significant distance from the dot).

Time spent scanning out-the-window in the undefined-tolerance condition was also similar to the larger conformance-band condition. The larger tolerance band supports better predictability of aircraft location along the taxi route than the undefined-conformance format. Pilots' ratings of the acceptability of eyes-in time, and the safety of taxiing with the display format, were both lower, although not significant, in the undefined-tolerance display condition than in the larger +/-405 ft tolerance band.

Percent time scanning out-the-window decreased when pilots received five speed changes as compared to two. However, no effect of speed changes was seen in the pilots' assessment of eyes-in time or workload ratings. More research is required to specify the acceptable maximum number of speed changes, or the acceptable minimum distance between changes.

No effect of 4DT speed ('slow' vs. 'fast') was evident in the conformance, eyes-out time, or subject safety/workload ratings. However, pilots indicated that it may be challenging to maintain slower speeds (e.g., 8 or 9 kts) in an actual aircraft, and may require more control inputs (e.g., braking) to do so. Likewise, pilots reported that they would be unlikely to maintain faster taxi speeds (e.g., 21–25 kts) while approaching a turn or the departure queue area, and therefore would increase brake use. Since excessive braking can affect the safety of takeoff (in case of an aborted takeoff), pilots use their brakes sparingly during taxi. More research needs to be conducted to understand the interactions among 4DT speed requirements, aircraft operating envelopes, and airport layout.

ACKNOWLEDGMENT

This work was funded by NASA's Airspace Operations and Safety Program (AOSP) / Airspace Technology Demonstration Project (ATD-2, Integrated Arrival / Departure / Surface) / Exploratory Development effort. The authors are indebted to Glenn Meyer (NTT DATA Services Federal Government, Inc.) for experimental and analysis software support and Nancy Haan for support throughout the experiment. The authors thank Nikolai Okuniek, Ingrid Gerdes, and the DLR team for their help in characterizing the TRACC outputs and taxi routes used to inform the experimental design and scenario development of this study.

REFERENCES

[1] Joint Planning and Development Office (JPDO). (2011). *Concept of Operations for the Next Generation Air Transport System, v3.2.* http://www.dtic.mil/dtic/tr/fulltext/u2/a535795.pdf

[2] Hooey, B. L., Cheng, V. H. L., Foyle, D. C. (2014). *A concept of operations for far-term Surface Trajectory-Based Operations (STBO).* (NASA TM–2014–218354). Moffett Field, CA: NASA ARC. https://hsi.arc.nasa.gov/groups/HCSL/publications/STBO%20ConOps_TM_2014_218354.pdf

[3] Jung, Y., Hoang, T., Montoya, J., Gupta, G., Malik, W., Tobias, L., & Wang, H. (2011). Performance Evaluation of a Surface Traffic Management Tool for Dallas/Fort Worth International Airport. *Ninth USA/Europe Air Traffic Management Research and Development Seminar, ATM2011 (Paper 92), Berlin, Germany, June 14–17, 2011.* http://www.aviationsystemsdivision.arc.nasa.gov/publications/2011/ATM2011_Jung.pdf

[4] National Aeronautics and Space Administration (NASA). (2017). *Airspace Technology Demonstration 2 (ATD-2).* https://www.aviationsystemsdivision.arc.nasa.gov/research/tactical/atd2.shtml

[5] Simmons, M. (2012). A functional analysis of integrated arrival, departure, and surface operations in NextGen. *Proceedings of the 31st Digital Avionics Systems Conference (DASC), Williamsburg, VA, October 14–19, 2012.*

[6] Federal Aviation Administration (FAA) Air Traffic Organization (ATO), Surface Operations Office. (2012). *U.S. Airport Surface Collaborative Decision Making (CDM) Concept of Operations in the Near-Term.* https://faaco.faa.gov/index.cfm/attachment/download/33926

[7] Eurocontrol. (2017). *Airport Collaborative Decision Making (A-CDM).*

[8] Audenaerd, L. F., Burr, C. S., & Morgan, C. E. (2009). *Surface Trajectory-Based Operations (STBO) Mid-term Concept of Operations Overview and Scenarios.* The MITRE Corporation, McLean, VA.

[9] Stelzer, E. K., Morgan, C. E., McGarry, K. A., Klein, K. A., & Kerns, K. (2011). Human-in-the-loop simulations of surface trajectory-based operations. *Ninth USA/Europe Air Traffic Management Research and Development Seminar, ATM2011, Berlin, Germany, June 14–17, 2011.* https://www.mitre.org/sites/default/files/pdf/11_0190.pdf

[10] Okuniek, N., Gerdes, I., Jakobi, J., Ludwig, T., Hooey, B. L., Foyle, D. C., Jung, Y. C., & Zhu, Z. (2016). A concept of operations for trajectory-based taxi operations. *Proceedings of the 16th AIAA Aviation Technology, Integration, and Operations Conference, Paper AIAA-2016-3753, Washington, DC, June 13–17, 2016.* https://hsi.arc.nasa.gov/groups/HCSL/publications/AIAA_ATIO_2016_NASA_DLR_4DT_ConOps.pdf

[11] Foyle, D. C., Hooey, B. L., Bakowski, D. L., Williams, J. L., & Kunkle, C. L. (2011). Flight deck surface trajectory-based operations (STBO): Simulation results and ConOps implications. *Ninth USA/Europe Air Traffic Management Research and Development Seminar, ATM2011 (Paper 132), Berlin, Germany, June 14–17, 2011.* http://hsi.arc.nasa.gov/groups/HCSL/publications/Foyle_ATM2011_4_15_11_finalpaper_web.pdf

[12] Gerdes, I. & Temme, A. (2012). Taxi routing for aircraft: Creation and Controlling – Ground movements with time constraints. *Second SESAR Innovation Days, November 27–29, 2012.* https://www.sesarju.eu/sites/default/files/SID_2012-05.pdf

[13] Shelton, K. J., Prinzel III, L. J., Arthur III, 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. In J. Guell and M. U. de Haag (Eds.), *Proceedings of SPIE 7328, Enhanced and Synthetic Vision 2009, Vol. 73280C.*

[14] Jones, D. R., Prinzel, L. J., Bailey, R. E., Arthur, J. J., & Barnes, J. R. (2014). Effect of traffic position accuracy for conducting safe airport surface operations. *Proceedings of the 33rd Digital Avionics Systems Conference (DASC), Colorado Springs, CO, October 5–9, 2014.*

[15] Bakowski, D. L., Hooey, B. L., Foyle, D. C., & Wolter, C. A. (2015). NextGen Surface Trajectory-Based Operations (STBO): Evaluating Conformance to 4-Dimensional Trajectories (4DT). In T. Ahram et al. (Eds.), *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, Procedia Manufacturing Vol. 3,* (pp. 2458–2565). Elsevier Procedia. https://hsi.arc.nasa.gov/groups/HCSL/publications/AHFE15_Bakowski_Hooey_Foyle_Wolter.pdf