Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)

Flight deck surface trajectory-based operations (STBO): Simulation results and ConOps implications

David C. Foyle
Human Systems Integration Division
NASA Ames Research Center
Moffett Field, CA 94035
David.C.Foyle@nasa.gov

Becky L. Hooey, Deborah L. Bakowski,
Jennifer L. Williams, Christina L. Kunkle
San Jose State University Research Foundation at
NASA Ames Research Center
Moffett Field, CA 94035

***Abstract* — The results of four piloted medium-fidelity simulations investigating flight deck surface trajectory-based operations (STBO) will be reviewed. In these flight deck STBO simulations, commercial transport pilots were given taxi clearances with time and/or speed components and required to taxi to the departing runway or an intermediate traffic intersection. Under a variety of concept of operations (ConOps) and flight deck information conditions, pilots' ability to taxi in compliance with the required time of arrival (RTA) at the designated airport location was measured. ConOps and flight deck information conditions explored included: Availability of taxi clearance speed and elapsed time information; Intermediate RTAs at intermediate time constraint points (e.g., intersection traffic flow points); STBO taxi clearances via ATC voice speed commands or datalink; and, Availability of flight deck display algorithms to reduce STBO RTA error. The results of these simulations show that when pilots are provided with STBO speed-only taxi clearances by ATC, pilots either have poor RTA compliance with acceptable workload and safety estimates, or have good RTA compliance with unacceptable workload and safety estimates. The presence of a flight deck error-nulling algorithm display allows pilots to comply accurately with STBO taxi RTA clearances while maintaining safety under acceptable workload. The need for flight deck capabilities (integrated avionics or an electronic flight bag, EFB, at a minimum) to conduct effective STBO taxi clearance operations and implications for the development of surface traffic management (STM) systems are discussed.**

Keywords-STBO; surface operations; NextGen; taxi; displays

I. INTRODUCTION

On a global basis, research is underway to design the next-generation airspace systems of the world. The SESAR [1] and EMMA2 [2] efforts in Europe and the NextGen [3] efforts in the United States are core programs of these new technology efforts. Under these programs, all phases of flight are being investigated: Pre-flight, taxi, take-off; Departure, Climb; Enroute cruise; Descent, approach; and, Landing, taxi, arrival. The present studies investigated the taxi-out departure environment (from the ramp area to the runway) in the NextGen environment.

Interviews with individual pilots and focus groups have informed these authors that current-day taxi-out and departure are the busiest phase of flight for the flight deck crew. During taxi-out, crew taxi operations include: Communicating with air traffic control (ATC) regarding the clearance; Maneuvering the aircraft; Navigating the taxi clearance by referring to airport signage and the airport taxi chart; and, Maintaining separation from other aircraft. In addition to these taxiing duties, the pilots also conduct duties associated with departure, including: Configuring the aircraft for flight; Verifying the flight plan and departure clearance information in the Flight Management System (FMS); Confirming final passenger and baggage weight loads; Communicating with cabin crew and passengers; and, Completing pre-departure briefings related to the normal departure and potential safety backup procedures in the case of such off-nominal events as an engine-out on take-off.

A. NextGen Surface Traffic Management Systems

Current NASA research efforts are aimed at the development of surface traffic management (STM) systems for ATC to provide optimized taxi clearances enabling efficient airport traffic operations and improving throughput. One such example is SARDA (Spot and Runway Departure Advisor) [4], which estimates spot-to-runway taxi times based on the current average. These average aircraft taxi times are then used to provide timing guidance to the ground controller to estimate when aircraft are likely to reach the departure runway.

In addition to having specified arrival times at departing runways, future STM systems are expected to also eliminate active runway crossing delays, and enable more efficient use of runways. These future STM systems will have associated aircraft arrival times at active runway thresholds so that aircraft can cross with minimum, or no delay, and at intermediate taxiway traffic flow points, enabling aircraft departure queue sequencing. Such future "full capability" STM systems will require aircraft to reach specified airport "time constraint points" with relatively precise timing. To achieve this precise timing, future versions of STM systems will likely require the aircraft to be an "active participant" in reaching specified locations at specified times. Future taxi clearances would have a speed- or time-based component to which the pilot must comply – these NextGen taxi operations have been referred to as "4D taxi" (with the 4th dimension referring to the time component), or surface trajectory-based operations (STBO).

This work was funded by the NASA Airspace Systems Program / NextGen Concepts and Technology Development (CTD) Project / Safe and Efficient Surface Operations (SESO) Element.

STM systems are envisioned to use dynamic algorithms to generate speed- or time-based taxi clearances for aircraft to calculate the most efficient movement of all surface traffic and enable precise surface coordination [5, 6]. Research [7] has shown that aircraft take twice as long to cross active runways when starting from a standstill compared to crossing without having to stop. Additionally, if pilots can reach a runway crossing or an airport intersection traffic flow merge point within a specific window of time that allows them to proceed without stopping and holding, this would result in shorter taxi times, increased fuel efficiency, and increased traffic throughput of the airport [7].

There are many variants of the 4D taxi/STBO concept related to the STBO taxi clearance as shown in Fig. 1. In the extreme STBO case ("Full STBO" in Fig. 1, right), the aircraft is required to follow a fully defined time/location profile, such that every point along the taxi clearance has a defined predictable time associated with it. A simpler, less extreme STBO case is when only one or two airport surface locations along the taxi clearance have required times of arrival (RTAs) (Fig. 1, left). For example, release from the airport "spot" and arrival at the departure runway end may be the only two points in the taxi clearance that have defined times. In some manner, the STM system provides speed or time commands to pilots, requiring arrival at intermediate time constraint points, certain determined airport "traffic flow points" (e.g., traffic merge points, active runway crossings, etc.) at specific times. The aircraft's speed or time may need to be adjusted if the pilot is unable to conform to the STBO command, or if other traffic is unable to comply creating a reduction in separation, or to meet other needs of the dynamic airport surface (e.g., runway crossings, etc.). Since the STBO taxi concept is in its infancy, current efforts aim to impact the design of the underlying ATC STM algorithms, so that the resulting STM system does not exceed pilot/aircraft performance capabilities. That is, to ensure that the STM system does not require the pilot to accept a time-based taxi clearance that cannot be safely executed.

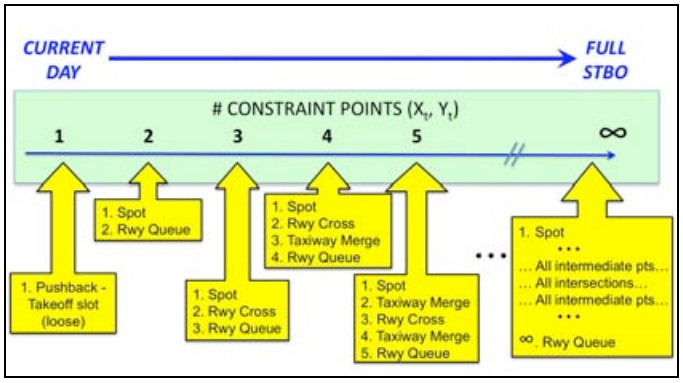

Figure 1. Variants of the STBO concept related to the STBO taxi clearance. STBO taxi clearances may have one or two (left) or an infinite number (right) of time "constraint points" – locations along the taxi clearance route that have an associated required time of arrival (RTA).

Currently there is no accepted STBO RTA design precision requirement. Under some conditions current-day operations can average 40 operations per runway per hour, that is, every 90 sec on average, and as close as 60 sec apart [5]. In the NextGen environment, in order to achieve improved efficiencies from current-day operations, it is reasonable to conclude that departure runway RTAs and runway crossing RTAs will need to be well within the observed 60-sec current-day operation window – possibly within a 30-sec window (i.e., RTA +/- 15 sec). It should be noted that fast-time simulation system studies are needed to determine this level of aircraft RTA precision or predictability necessary to enable various detailed STBO concepts.

As a minimum requirement, future implementations of STBO 4D taxi clearances will require new information on the flight deck (i.e., coordinated time information, airport taxiway distances, etc.) and may likely require advanced displays to support pilots during taxi operations. Defining that information and determining its impact on flight deck and surface operations is the goal of the research presented in this paper.

II. Experiment 1: Speed vs. Time Information

A. Experiment Objective

As an intial study, it is necessary to determine the relative contribution of flight deck speed information and time information on pilot STBO 4D taxi performance. In addition, as a sensitivity analysis, taxi route distance and required speed are manipulated. In Experiment 1, pilots are required to complete taxi routes while achieving specific average speeds or completion times when speed, time or both speed and time information are available.

B. Method

1) Participants: Eighteen current commercial airline Captains or First Officers, with a minimum of 500 hours as Captains within the last 4 years, participated in the experiment. Mean age was 49 yrs; mean flight hours as captain was 8,383 hrs. All pilots were male.

2) Flight Simulation: The study was conducted in a medium-fidelity part-task simulator in the Human-Centered Systems Laboratory (HCSL) at the NASA Ames Research Center with Boeing 737 modeled dynamics. The simulated environment modeled Dallas/Fort Worth International Airport with 1,200 ft visibility. The forward out-the-window scene was rear projected with visual angles: 53.1 deg (Horizontal, H) by 41.1 deg (Vertical, V). Side window views were displayed on two side monitors 31.9 deg (H) by 24.2 deg (V). A static, north-up airport chart and text taxi clearance were displayed on a 18.5 deg (H) by 13.5 deg (V) monitor in front of the pilot below the out-the-window scene. Pilots controlled the simulated aircraft using a tiller, throttle, and rudder toe-brakes. While taxiing, pilots wore an Applied Science Laboratory Model 501 head-mounted eyetracker.

The current instantaneous ground speed indicator, elapsed time, required taxi speed, and RTA were superimposed graphically over the out-the-window screen simulating a Head-Up Display (HUD). For the Speed format, current ground speed (Fig. 2, boxed, upper left) was displayed with commanded speed below it. For the Time format, elapsed time

(Fig. 2, boxed, upper right) was displayed with required time of arrival (RTA) below it. Each display item measured 1.2 deg (H) by 0.6 deg (V). For the Speed/Time format, current ground speed, commanded ground speed, elapsed time, and commanded time were displayed simultaneously. A text triad (Fig. 2, middle) indicated the ownship's current taxiway and the upcoming left and right taxiways.

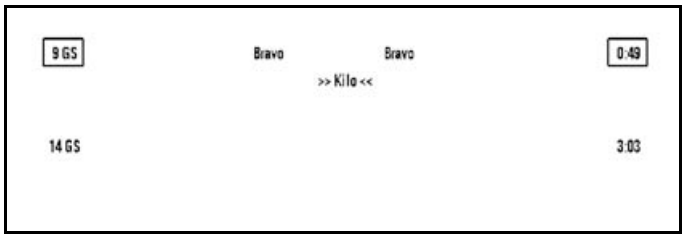

Figure 2. Speed and time information presented superimposed over the sky in the out-the-window airport view. (See text for description).

A static north-up diagram of the airport and taxi clearance text were available during all trials. At the beginning of each trial, during clearance reading/readback, the map displayed an ownship icon at the starting point of the clearance and a magenta line of the cleared taxi route. After clearance readback by the pilot to ground control, the magenta route and ownship icon were removed, but the airport map and text taxi clearance remained during the trial.

3) Experimental Design: The experiment was a 3 (format) x 3 (distance) x 4 (required speed) within-subjects design. The three levels of command format were speed, time, and both speed and time. For the speed format, pilots were instructed to taxi to a runway crossing at a commanded average speed. With the time format, pilots were instructed to taxi to a specific runway crossing arriving at a commanded RTA. In the speed/time format, pilots were required to comply with both a required speed and its corresponding commanded RTA. The commanded speeds were based on combinations of the four required speeds (10, 14, 18 and 20 kts) and three route distances (3,000, 6,000, and 12,000 ft).

The experiment was comprised of three blocks of 15 trials each (3 practice and 12 experimental trials). Each block consisted of one of the three command formats. Block presentation order was counterbalanced across subjects, and the commanded speeds and route distances were randomized within each block and matched for each format.

4) Procedure: Pilots completed a series of initial tasks before the experimental trials. In three familiarization trials pilots were given a taxi clearance and a ground speed indicator and were to taxi at a variety of comfortable speeds. They then completed two or three training trials and were required to taxi at 10, 20, and 30 kts without a speed display. After indicating that they reached the target speed, their actual speed was displayed. This was repeated until they had twice taxied within 20% of the target speed (thus ensuring that they were appropriately calibrated to the simulator). Finally, pilots taxied one practice trial for each command format.

For each trial, the participants received a verbal taxi clearance from ground control, a text taxi clearance at the bottom of the airport diagram, and a taxi command in the speed, time, or speed/time format. Ground control read the taxi clearance and confirmed readback. Pilots were to complete the route with the required average speed, or arrive at the commanded time, as closely as possible. Some realistic traffic was present, but taxi routes were conflict free. After each trial, pilots were shown their actual average speed or RTA, as well as the commanded speed, commanded RTA, or both (depending on condition). Total testing lasted about 8 hrs.

C. Results and Discussion

RTA absolute error (the absolute value of the difference between the commanded RTA and pilots' actual arrival time at the designated runway crossing) assesses the accuracy of compliance with the STBO commands. A 3 (format) x 3 (distance) x 4 (required speed) repeated measures ANOVA was conducted on RTA absolute error. There was a main effect of format ($F(1,21)=5.95$, $p<.05$) and a significant interaction between distance and required speed ($F(3,54)=4.00$, $p<.01$). RTA absolute error for the speed/time format was significantly lower than RTA absolute error for the speed format ($t(17)=3.72$, $p<.01$); and the time format ($t(17)=3.43$, $p<.01$). Pilots' RTA conformance was more accurate with the speed/time format than with either the speed or time formats (see Fig. 3).

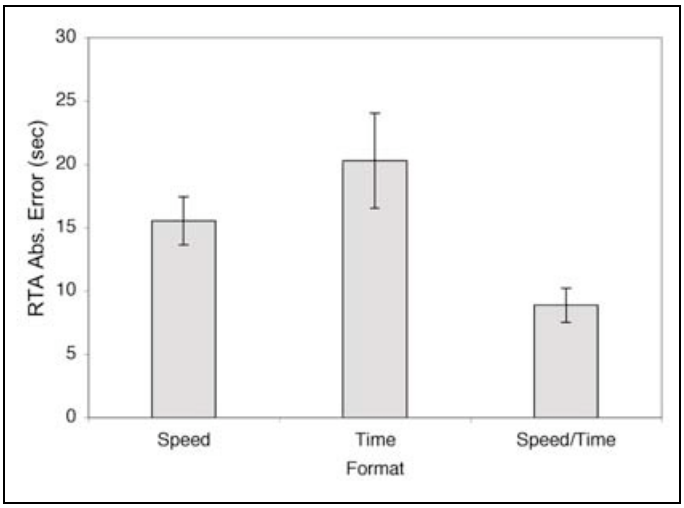

Figure 3. Mean required time of arrival (RTA) absolute error for display formats. Error bars = +/- 1 standard error.

RTA error was measured as the signed difference between the commanded RTA and the pilots' actual arrival time, calculated by subtracting the commanded RTA from the actual arrival time (i.e., negative values of RTA error indicate early arrival, and positive indicate late arrival). A 3 (format) x 3 (distance) x 4 (required speed) repeated measures ANOVA was conducted on RTA error. There were no main effect or interactions of format. There was a significant interaction between distance and required speed ($F(5,78)=7.22$, $p<.001$). Simple effects analyses of the distance by required speed interaction showed that RTA error for required speed differed for 3,000 ft ($F(3,71)=13.19$, $p<.001$); 6,000 ft ($F(3,71)=22.43$, $p<.001$); and 12,000 ft ($F(3,71)=14.26$, $p<.001$) routes. Simple

effects analyses also revealed that RTA error differed between route distances for the required speeds of 10 kts ($F(2,53)=4.65$, $p<.05$); 14 kts ($F(2,53)=8.33$, $p<.01$); and 18 kts ($F(2,53)=7.23$, $p<.01$); but not for 22 kts. As seen in Fig. 4, with the exception of the 22 kts required speed, pilots arrived at the runway crossings earlier as distance increased, and across distances pilots arrived earlier as required speed decreased.

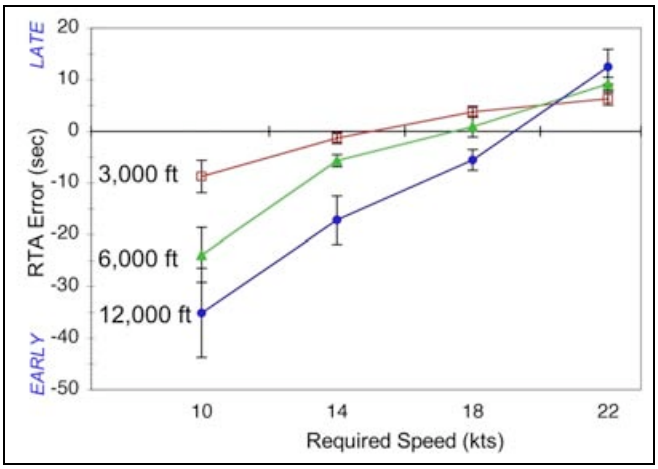

Figure 4. Mean required time of arrival (RTA) error by distance and required speed, averaged across display format. Error bars = +/- 1 standard error.

Percentage of dwell time was calculated as the proportion of dwell time that pilots were looking at various pieces of display information. Due to technical difficulties, eye-tracking data was not collected for one pilot. A 3 (format) x 3 (distance) x 4 (required speed) repeated measures ANOVA was conducted on the percentage of dwell time on display information. There was a significant main effect of format ($F(2,32)=14.88$, $p<.001$. As seen in Fig. 5, pair-wise comparisons of format indicated that the percentage of dwell time was significantly lower for the time format than for the speed ($t(16)=5.10$, $p<.001$); and the speed/time formats ($t(16)=4.48$, $p<.01$).

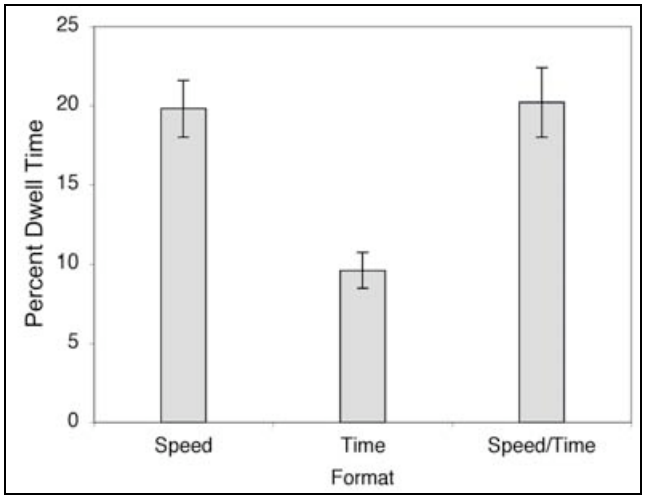

Figure 5. Mean percentage of eye dwell time for the three formats. Error bars = +/- 1 standard error.

In order to understand the time course of information use, routes were divided into quartiles. A 3 (format) x 3 (distance) x 4 (required speed) x 4 (quartile) repeated measures ANOVA was conducted on percentage of dwell time, revealing a significant interaction between format and quartile ($F(3,46)=23.80$, $p<.01$). Pilots viewed the speed format display most at the beginning of the route and looked at the time format display most near the end of the route, whereas with the speed/time format the quartile pattern is a combination of the speed and time format patterns (see Fig. 6). (Further analysis revealed that in the speed/time format, pilots mostly used speed information in quartiles 1 - 3, and transitioned to mostly using time information in the final quartile of the route).

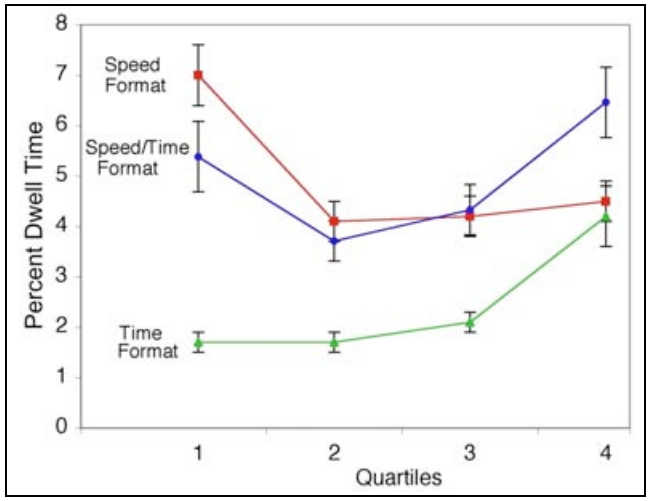

Figure 6. Mean percentage of eye dwell time on speed and/or time information by format and route quartile (1 = First quarter of the route, 4 = Last quarter of the route). Error bars = +/- 1 standard error.

In summary, results indicate that pilots' RTA error was the lowest with the speed/time format, but when analyzed by route quartile, performance with the speed format was equivalent to the speed/time format during the first three quartiles and did not diverge until the ends of the routes. Pilots also taxied more accurately with shorter routes and faster required speeds. RTA error data revealed that pilots arrived equally early regardless of format, but arrived earlier with slower required speeds and longer distances. Finally, percentage of dwell time data indicate that pilots looked at the speed and speed/time displays more than the time display. Additionally, quartile analyses revealed that pilots' patterns of display usage when given both speed and time information were equivalent to the patterns of usage with the individual speed and time formats in that they viewed/used speed information primarily at the beginning of the route and viewed/used time information at the end of the route.

III. Experiment 2: Commanded Speed – Without Speed Profiles or Conformance

A. Experiment Objective

The simplest STBO 4D taxi clearance, in terms of pilot procedures and flight deck equipage, is for ATC to provide a required speed to individual aircraft. This might be supplied to the Ground Controller by tower-based decision support tool

automation. Because of the distance effect (larger RTA error with longer routes) observed in Experiment 1, Experiment 2 replicates Experiment 1 with the addition of intermediate time constraint points (and new speed requirement) and higher fidelity flight deck displays. In Experiment 2, Ground Control provided the pilot a taxi clearance with a required speed to maintain. Pilots in this experiment were not required to follow any specific acceleration/deceleration speed profiles.

B. Method

1) Participants: Eight commercial pilots (6 Captains, 2 First Officers), current or recently retired, participated in the study. Mean pilot age was 49, and mean flight hours logged was 5,029 hrs. One pilot was female, 7 were male.

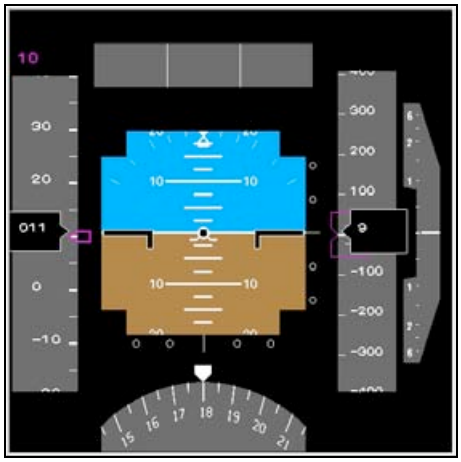

Figure 7. Primary Flight Display (PFD).

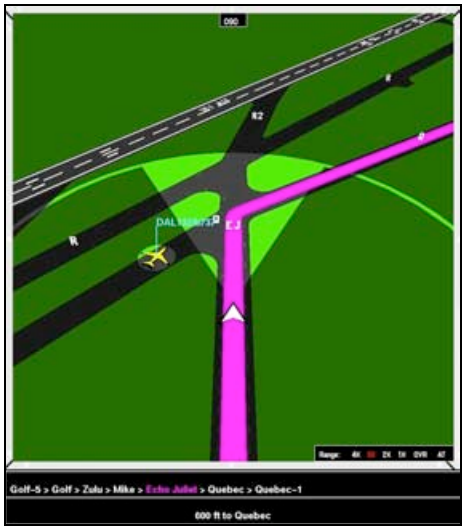

Figure 8. Taxi Navigation Display (TND).

2) Flight Simulation: The study was conducted in the same simulator as in Experiment 1. The airport environment was the Dallas/Fort Worth Airport (DFW) with high visibility and distant fog/haze conditions. In addition, the simulator flight deck included a Primary Flight Display (PFD), Navigation Display (ND), Taxi Navigation Display (TND), Datalink Display, and an Electronic Checklist.

The PFD was modified for taxi operations – the speed scale (see Fig. 7, left) was re-scaled from 0-60 kts to support taxi operations. The commanded ground speed (e.g., 10 kts in Fig. 7, top left) was displayed both digitally in magenta directly above the speed tape and as a magenta analog pointer ("speed bug"). Current speed (e.g., 11 kts in Fig. 7) was shown as a sliding indicator with digital value inside.

To assist in airport navigation, the simulator flight deck included a dynamic TND that depicted the airport layout. At the start of each trial the TND showed the entire airport in north-up view to support route planning. It changed to track-up perspective mode (Fig. 8) when the pilot started taxiing. The ownship aircraft's position (white chevron), and other aircraft traffic were updated in real time. The taxi clearance, presented graphically as a magenta route and as text below the TND, indicated the cleared route with positive cleared-to-cross runway clearance.

3) Experimental Design: The experiment was a within-subjects design with two factors, Number of time constraint points (1, 3, or 5) and Commanded speed (10, 14, 18, or 22 kts). These two factors were crossed factorially and assigned randomly to 12 unique taxi routes. These 12 taxi clearance routes were repeated twice during the testing day with different airport traffic configurations, yielding 24 trials (3 flow point values x 4 speeds x 2 repetitions).

4) Procedure: Pilots completed departure taxi scenarios from a ramp departure spot to departure runways (mean taxi time = 9.5 min). At the beginning of each trial, the Ground Controller/Experimenter issued a verbal taxi clearance to the departing runway that also appeared in text and graphically on the TND. After the pilot reviewed the map and clearance, the TND switched to the track-up perspective view (see Fig. 8), and the trial began. Data were collected after the pilot had completed 4 such familiarization taxi departure trials.

An auditory chime and the verbal cue “change speed” accompanied each taxi segment transition at the intermediate time constraint point location. While taxiing, pilots received a datalinked departure clearance for verification, and were to monitor the status of an electronic checklist. These secondary tasks are not described here; see [8] for details.

C. Results and Discussion

The primary measure of pilot performance on the taxi task was RTA error, calculated by subtracting the RTA from the arrival time. Pilots did not receive an explicit commanded RTA – they received a commanded speed that they were required to follow on straight segments, told not to exceed 15 kts in turns, and instructed to accelerate/decelerate “aggressively”. For analysis, RTA (and thus RTA error) was calculated using the taxi route segment length and the ATC-commanded speed for the straight segments, with an assumed underlying speed profile of 2 kts/sec acceleration/deceleration; turn speeds were 15 kts for commanded speeds of 18 and 22 kts, and equal to the straight segment speed for commanded speeds of 10 and 14 kts.

Positive RTA errors indicate that the pilot taxied too slowly and therefore arrived late. Negative RTA errors indicated that

the pilot taxied too quickly and therefore arrived early. A 3 (number of time constraint points) by 4 (commanded taxi speeds) within-subjects ANOVA (see Fig. 9) showed that there was an interaction between number of time constraint points and commanded taxi speed ($F(6,42)=6.79$, $p<.001$) and a main effect of commanded taxi speed ($F(3,21)=24.87$, $p<.001$).

In this experiment, pilots were to follow the taxi route clearance with the associated ATC speed command. With speed commands only, pilots exhibited more difficulty maintaining a relatively fast taxi speed (18 or 22 kts) for a long distance, as in the 1 time constraint point condition – but RTA error was reduced for these speeds by adding 3 or 5 time constraint points. Pilots performed best at a taxi speed of 14 kts, with negligible RTA error regardless of the number of time constraint points. Presumably, it was easier for pilots to maintain an average taxi speed of 14 kts, because they did not need to slow down for turns and then resultantly correct for the slower speed. With a commanded speed of 10 kts, there was some reduction in RTA error with the addition of multiple (3 or 5) time constraint points, which served to decrease the distance of each segment. (Presumably, this is due to the same "route distance effect" seen in the data of Experiment 1.)

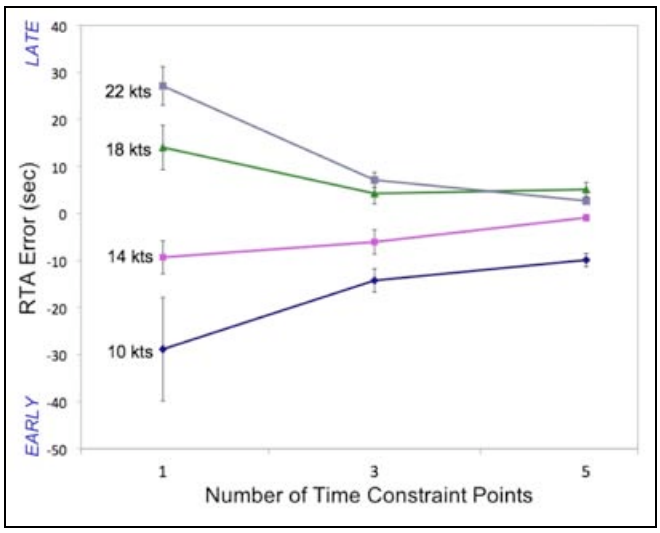

Figure 9. Mean required time of arrival (RTA) error as a function of commanded speed and number of time constraint points. Error bars = +/- 1 standard error.

IV. Experiment 3: Commanded Speed with Speed Profiles

A. Experiment Objective

Following up on the previous results, Experiment 3 was aimed at determining if RTA error would be reduced if pilots were explicitly required to follow the speed acceleration/deceleration profile used to calculate the route RTA. Pilots in Experiment 3 were again required to follow taxi commanded straight and turn speeds, but in contrast to the previous experiment, specific acceleration/deceleration speed profiles were required to be followed. In addition, the impact of a bounding range around the commanded speed was assessed.

B. Method

1) Participants: Eighteen commercial pilots (13 Captains, 5 First Officers), current or recently retired, participated in the study. Mean pilot age was 45 years and mean flight hours logged was 3,832 hrs. One pilot was female, 17 were male.

2) Flight Simulation: The study used the same HCSL simulator and visibility conditions (high visibility and distant fog/haze) at DFW airport as in Experiment 2. For this experiment, the two left/right side monitors subtended 29.6 deg visual angle. The same PFD (except that no commanded speed was displayed) and TND configuration as in Experiment 2 were used.

3) Experimental Design: The experiment consisted of three within-participant factors, Speed-Conformance implementation (Undefined and Defined), Number of time constraint points (1, 3, or 5) and Commanded speed (14, 18, or 22 kts). The three experimental factors were crossed factorially to create nine nominal trials in each of the two speed conformance conditions.

In the Undefined Speed Conformance condition, pilots were instructed to taxi as close to the verbal commanded speed as was reasonable. No required speed conformance range or performance feedback was provided in this condition. However, in the Defined Speed Conformance condition, pilots were instructed to taxi within +/-1.5 kts of the commanded speed. When ground speed exceeded the +/-1.5 kt range for more than a continuous 5-sec period, ATC delivered an automated verbal alert, "NASA227, check speed", repeating every 10 sec until the pilot's speed returned to within the +/-1.5 kt range. The ATC "check speed" alert was disabled immediately after a speed command and near turns. For all pilots, the Undefined Speed Conformance condition was tested first, so that performance represented the pilots' "natural" uninstructed speed conformance level, followed by the Defined Speed Conformance trials. After testing the Undefined Speed Conformance condition, and prior to testing the Defined Speed Conformance condition, pilots completed a single Baseline Current Day trial. In the Baseline Current Day trial, pilots were not given a commanded speed and were instructed to taxi "as they would normally in the real world". In addition to the nine nominal trials, three off-nominal trials (not discussed here) were tested.

4) Procedure: Pilots completed departure taxi scenarios from a ramp departure spot to a departure runway. A verbal ATC command providing a changed taxi speed (e.g., "NASA 227, taxi at 14 kts") accompanied each taxi segment transition (14, 18, or 22 kts) at the time constraint point. Segment distances and speed changes were not depicted on the TND. In addition to the speed command that pilots received at the start of each segment, two specific aircraft control/speed profile instructions were given to pilots, and applied to all trials in both implementations (with the exception of the baseline trial). These aircraft control/speed profile instructions were: 1) Taxi all turns at 14 kts; and, 2) Accelerate/decelerate at 2 kts/sec

(e.g., 7 sec acceleration for 0 kts to 14 kts; 2 sec for a speed change from 22 kts to 18 kts, etc.).

Required time of arrival (RTA) for each segment, the primary measure of pilot performance, was calculated using the length of the segment, the number of turns within the segment, nominal turn speed (14 kts), commanded speed, and nominal acceleration/deceleration rates. The final segment in each trial ended several hundred feet before the threshold of the departure runway. Trial average length was 8 min 42 sec. Secondary tasks (not discussed here) of electronic checklists and departure clearance verification were required during taxi.

C. Results and Discussion

1) Taxi Required Time of Arrival (RTA) Error: The taxi RTA error analyses included nine nominal trials in each speed conformance range condition (Undefined and Defined). The Current Day Baseline and three off-nominal trials were excluded from these analyses.

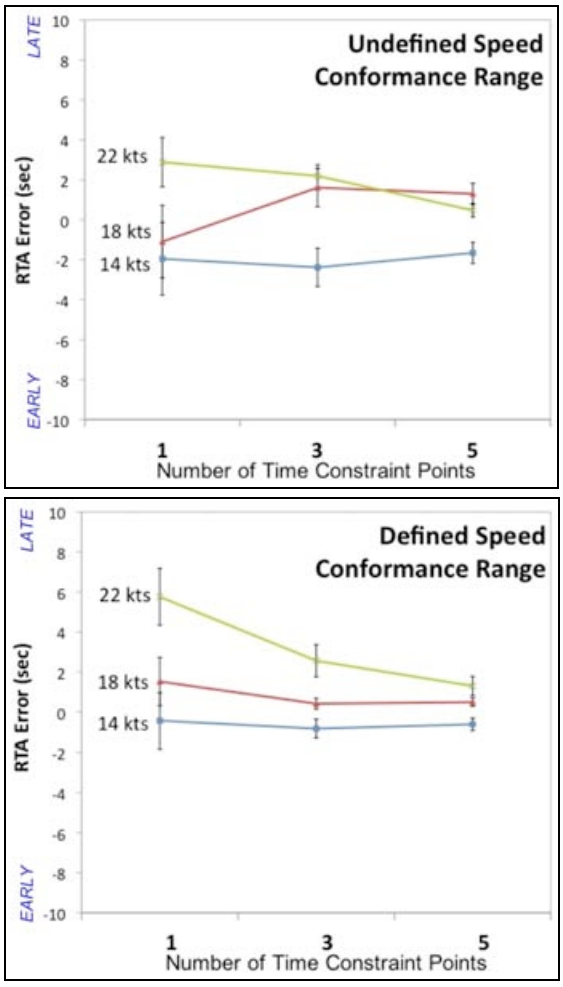

Figure 10. Mean required time of arrival (RTA) error as a function of commanded speed and number of time constraint points for the Undefined and Defined speed-conformance conditions. Error bars = +/- 1 standard error.

The primary measure of pilot performance on the taxi task was RTA error, calculated by subtracting the RTA from the observed arrival time. The RTA values were calculated using the nominal/instructed speed profile, considering the number of turns, nominal turn speed of 14 kts, and 2 kts/sec acceleration/deceleration before and after turns, for initial taxiing, and for commanded speed changes. RTA error provides an indication of the presence of an 'early' or 'late' bias. Positive RTA errors indicate that the pilot taxied too slowly and arrived late, negative RTA errors indicated that the pilot taxied too quickly and arrived early.

A 2 (speed-conformance implementation) by 3 (number of time constraint points) by 3 (commanded taxi speeds) within-subjects ANOVA indicated an interaction between number of traffic segments and speed ($F(4,68)=3.44$, $p=.013$), and an interaction between implementation and number of time constraint points ($F(2,34)=4.44$, $p=.019$). A main effect of speed was also found ($F(2,34)=21.83$, $p<.0001$). As seen in Fig. 10, RTA error is quite good in all conditions, and is much better than found in Experiment 2 where pilots were not required to follow specific speed and acceleration/deceleration profiles.

2) Percent Dwell Time on PFD Speed: The percentage of time that the pilots fixated (percent dwell time, PDT) on the current speed read-out displayed on the instrument panel PFD is shown in Fig. 11. Relative to the baseline current-day condition, pilots viewed the head-down PFD speed display 2.4 and 3.3 times more in the Undefined and Defined speed conformance conditions, respectively ($F(2,20)=41.29$, $p<.0001$).

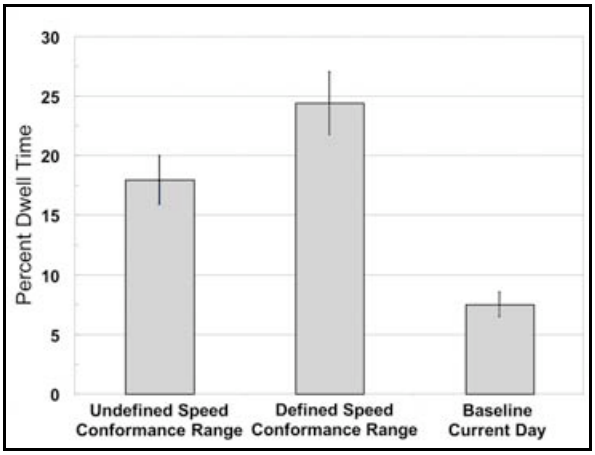

Figure 11. Mean percent dwell time on PFD speed for the three conditions (Undefined and Defined Speed Conformance bounds, Current Day Baseline). Error bars = +/- 1 standard error.

Compared to the current-day baseline condition of 7.55 percent dwell time, for the two speed conformance conditions, pilots spent 17.89 and 24.39 percent of the trial looking at the speed display. In absolute as well as relative terms, this is a large percentage of the trial to be looking head-down (or even head-up if on a HUD) at the speed display when the main duties of the taxiing Captain are to navigate and control the aircraft and maintain awareness and separation from other airport taxiing aircraft.

Data from this experiment indicated that when pilots are given STBO taxi clearances and are required to control their

aircraft according to precise acceleration/deceleration speed profiles, RTA conformance is quite good (on average 6 secs or better). However, associated with this good RTA conformance, pilots spend an inordinate amount of time viewing/tracking their speed. Given that taxiing relies so heavily on out-the-window airport navigation and aircraft/vehicle/pedestrian separation, the eyetracking data raises safety concern, to be addressed later in this paper.

V. Experiment 4: Error-Nulling Algorithm/Display

A. Experiment Objective

The previous 3 studies assessed pilot RTA error when given taxi clearances with speed and time requirements, and had available simple readout displays of current speed or elapsed time displays. Experiment 4 assessed more advanced equipage capabilities, such as advanced avionics or an electronic flight bag (EFB) display. In this study, pilots are provided a PFD speed display driven by an "error-nulling" algorithm that computes the necessary current speed to arrive at the required location (intersection or runway) at the RTA.

B. Method

1) Participants: Eight male commercial pilots (7 Captains, 1 First Officer), current or recently retired, participated in the study. Mean pilot age was 42 yrs; mean flight hours logged was 6,143 hrs.

2) Flight Simulation: The same physical setup, displays (with the exception of the PFD; see below) and general procedure were used as reported in Experiment 2.

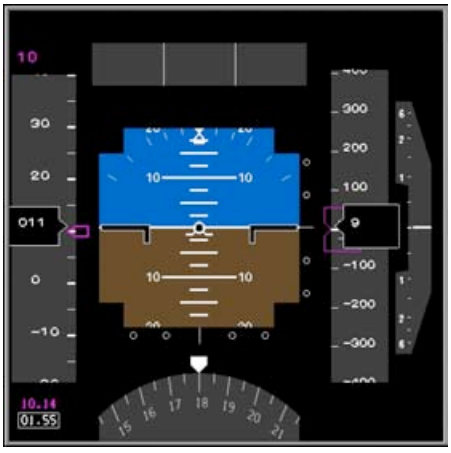

Figure 12. Prototype NextGen PFD. Speed tape (left) shows Commanded Speed = 10kts; Current Speed = 11kts. Elapsed taxi time = 1:55 min/sec; Total required taxi time = 10:14 min/sec.)

The PFD was identical to that used in Experiment 2, with the exception that time information was included (see Fig. 12). The lower left area of the PFD included Elapsed Time in a white box (in min.sec), counting upwards from zero, and RTA in magenta (in min.sec). An error-nulling algorithm dynamically compensated for speed-maintenance errors by adjusting the current commanded speed according to the remaining RTA and remaining distance to the time constraint point (according to the calculation: Current Commanded Speed = Remaining Distance / Remaining Time). Thus, by following the currently commanded speed, the aircraft would arrive at the time constraint point at the RTA. With the error-nulling algorithm, pilots received implicit performance feedback relative to the RTA. For example, if pilots were slow, the algorithm would increase the commanded speed, attempting to drive the pilots toward on-time RTAs (i.e., zero RTA error).

The TND was identical to that used in Experiment 2, with the exception that the time constraint point location was shown graphically on the TND as a yellow bar across the cleared route.

3) Experimental Design: The same factors and routes were used as in Experiment 2. The experiment was a within-subjects design with two factors, Number of time constraint points (1, 3, or 5) and Commanded speed (10, 14, 18, or 22 kts). Similarly, there were a total of 24 identical taxi trials (3 time constraint point values x 4 speeds x 2 repetitions).

4) Procedure: Pilots completed departure taxi scenarios from a ramp departure spot to a departure runway following the same general procedures as in Experiment 2. However, when reaching a time constraint point, the pilots received an auditory chime with the verbal cue, "checkpoint".

C. Results and Discussion

The primary measure of pilot performance on the taxi task was RTA error, calculated by subtracting the RTA from the arrival time. RTA error provides an indication of the presence of 'early' or 'late' bias. Positive RTA errors indicate that the pilot taxied too slowly and therefore arrived late. Negative RTA errors indicated that the pilot taxied too quickly and therefore arrived early. A 3 (number of time constraint points) by 4 (commanded taxi speeds) within-subjects ANOVA (see Fig. 13) revealed an interaction between number of time constraint points and commanded taxi speed ($F(6,42)=3.67$, $p=.005$). There were also main effects of commanded taxi speed ($F(3,21)=7.30$, $p=.002$) and number of time constraint points ($F(2,14)=4.86$, $p=.025$).

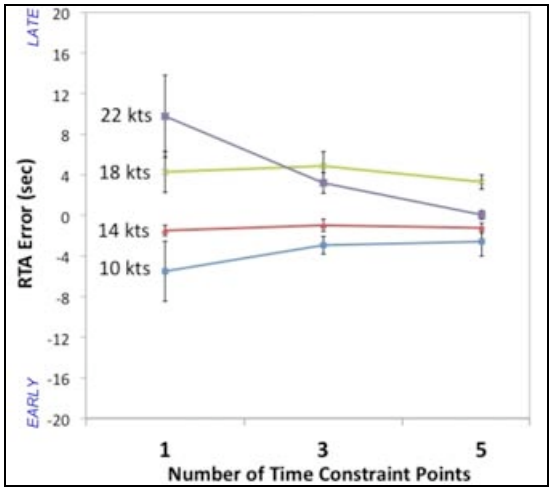

Figure 13. Mean required time of arrival (RTA) error as a function of commanded speed and number of time constraint points. Error bars = +/- 1 standard error.

As can be seen in Fig. 13, the availability of an error-nulling algorithm driving a commanded-speed display allowed pilots to reach the runway or time constraint point within approximately 10 sec of the RTA for all speeds and number of time constraint points tested (similar to the results of Experiment 3). Additionally, as will be discussed in the next section (and seen in Fig. 14), pilots reported that the use of the algorithm/display did not interfere with their attention to the airport environment. The nature of the algorithm is such that moment-by-moment attention to speed is not required. Momentary speed fluctuations have relatively little effect on speed since it is averaged over the remaining portion of the taxi route.

VI. Usage and Safety Implications

The taxi performance data presented earlier is not the whole story regarding flight deck STBO. New flight deck requirements can put new pressures on the crew, which may compromise or impact safety. The previously presented eye-tracking data in Experiment 3 indicate that when required to taxi their aircraft with a specific acceleration/deceleration speed profile, pilots spend a large percentage (18 - 24%) of time head-down viewing their ground speed, compared to what they would do under current day operations (8%).

Similarly, post-trial and post-study questions also inform potential safety issues. Also in Experiment 3, pilots were asked the following post-study question: "Would the demand of having to maintain the required speed conformance range compromise safety in the real world?" The data showed that 14 of the 18 pilots tested responded that the demand of having to maintain the required speed conformance range in the real world would compromise safety ($X^2(1)=5.56$, $p=.018$). Taken together, the eye-tracking data and the safety question raise safety concerns for conditions when pilots must precisely follow speed profiles.

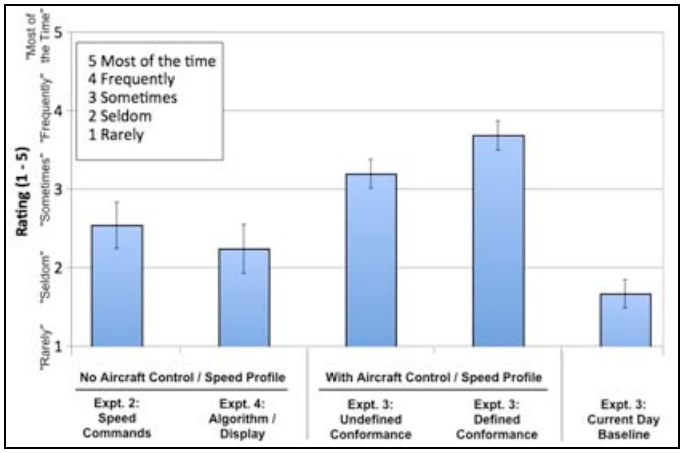

Figure 14. Mean rating response (see scale above) to post-trial question, "How often did you find yourself focusing on the PFD Speed or Time display, when you should have been paying attention to the external taxiway environment?" Error bars = +/- 1 standard error.

In Experiments 2, 3, and 4, pilots were asked to rate (1=Rarely, 5=Most of the Time; See Fig. 14) their response to the post-trial question, "How often did you find yourself focusing on the PFD Speed or Time display, when you should have been paying attention to the external taxiway environment?" As can be seen in Fig. 14, the results for Experiment 3, which required taxi according to a specific acceleration/deceleration profile, mean response values were in the range "Sometimes" - "Frequently". That is, pilots responded that "sometimes/frequently" they were not paying enough attention to the airport environment. Lower ratings were obtained in Experiments 2 ($t(24)=1.95$, $p=.064$, 1-tail significance) and 4 ($t(24)=2.78$, $p=.010$), which did not require the following of acceleration/deceleration speed profiles, with mean values in the range "Seldom" - "Sometimes". That is, pilots responded that only "seldom/sometimes" they were not paying attention to the airport environment. It should be noted that, as one would expect, the lowest ratings occurred in Experiment 3's current day baseline condition, with mean ratings of "rarely/seldom". These data are further evidence of the challenging nature of STBO taxiing with the requirement to follow a specified taxi speed acceleration/deceleration profile.

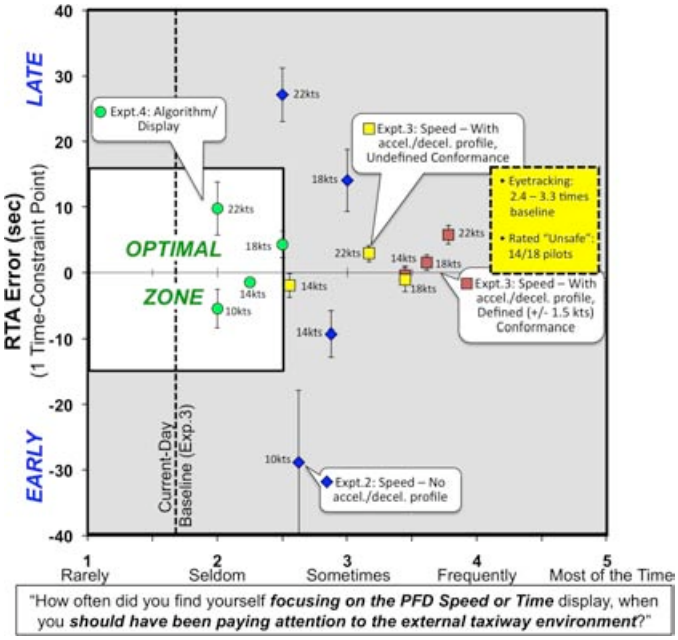

Figure 15. Mean ratings for attention question (from Fig. 14) by mean RTA errors for the commanded speeds tested in Experiments 2, 3, and 4. (Only data from the 1 time-constraint point [departing runway] conditions are shown; dashed line represents "current-day baseline" condition rating). The white "optimal zone" is the area with both good RTA conformance and with appropriate attentional allocation. Error bars = +/- 1 standard error.

The post-trial attention question ratings are shown in Fig. 15 as a function of the corresponding mean RTA errors for the speeds tested in Experiments 2, 3, and 4 (1 time-constraint point only, i.e., the departing runway). The white "optimal zone" is labeled as such because it represents the area in which mean RTA performance was both within a 30-sec RTA performance window (RTA +/- 15 sec) and yielded an appropriate rating regarding attentional allocation (i.e., a response of 2.5, the middle of the "Seldom-Sometimes" rating, or less). As can be seen, only the data from Experiment 4 (with the "error-nulling algorithm/display") lie within the "optimal zone", suggesting that only that STBO condition allows for good RTA conformance with appropriate attentional allocation.

The point here is not to draw absolute safety assessments from these data, but to understand the relative degree to which the various flight deck requirements may affect safety precursors (such as time looking out the window). Clearly, the present data suggest that STBO flight deck procedures that require precise tracking of acceleration/deceleration speed profiles yield increased safety concerns, whether measured with eye-tracking data or rating assessment.

VII. General Discussion and Summary

A series of four piloted medium-fidelity simulation experiments explored pilot taxi RTA performance in the NextGen STBO environment. The results of these experiments inform flight deck equipage and procedure requirements, as well as algorithm development of ATC/STM systems in the NextGen future airspace environment.

A. Flight Deck Implications

Pilot RTA conformance for STBO clearances, in the form of ATC taxi clearances with associated speed requirements, was found to be relatively poor, unless the pilot is required to follow a precise speed and acceleration/deceleration profile. However, following a precise speed profile results in inordinate head-down attention to tracking current ground speed, leading to potentially unsafe operations. Mitigating these results, and providing good taxi RTA performance without the associated safety issues, is a flight deck avionics or EFB solution such as the "error-nulling algorithm/display" described in Experiment 4. Such a "human-centered design" solution enables pilots to meet the taxi route RTA without moment-by-moment tracking of ground speed. Specifically, the proposed human-centered designed avionics or EFB "error-nulling algorithm/display" allows the pilot to view the STBO information when the pilot determines it is necessary and when workload allows, thus enabling the pilot to spread his/her attention appropriately and strategically on aircraft separation, airport navigation, and the many other flight deck tasks concurrently required. Clearly, a poorly designed display would not have these results.

We define a Human-Centered Design for a display or system, as one that:

- Is intuitive and natural
- Has readily accessible information
- Supports human capabilities (e.g., perceptual processing)
- Mitigates human limitations (e.g., memory failures)
- Enables appropriate task-usage strategies
- Has specific features supported by a design trace of human factors principles or empirical results

B. Surface Traffic Management (STM) System Implications

The results of these experiments have major implications for ATC/STM algorithm development. The data indicate a number of implications regarding specific parameters. Pilots have a tendency to arrive early with slow required speeds (i.e., 10 kts), and late with faster required speeds (i.e., 22 kts). This implies that ATC/STM algorithms should operate with middle-range speeds, similar to that of non-STBO taxi performance (see [9]). Route length has a related effect: Long taxi routes (i.e., 12,000 ft, typical for many airports) increase the earliness with slow speeds and the lateness with faster speeds. This is likely due to the "open-loop" nature of the task. That is, the speed error compounds over a longer time with longer routes. Results showed that this may be mitigated by imposing a small number of time constraint points each with their own RTAs. This has the resultant effect of turning a long route into a series of shorter routes – and thus improving RTA performance.

Most important is the impact that these data have for NextGen STM system ConOps development. The results of these experiments imply that without advanced display concepts, it is not reasonable to expect pilots to taxi under a "Full STBO" ConOps (as shown in Fig. 1) in which pilots are expected to be at a predictable (x,y) airport location for every time (t). An STBO ConOps with a small number of intermediate time constraint points and the departing runway, however, is feasible, but only with flight deck equipage similar to that used in Experiment 4 (the "error-nulling algorithm/display").

Acknowledgments

The authors are indebted to Glenn Meyer (Dell Services, Federal Government) for experimental and analysis software support, Martin Schwirzke (formerly with San Jose State University Foundation) for his help with the conduct of Experiments 2 and 4, and to Dr. Kevin Jordan (San Jose State University) for his data analysis expertise. Portions of these experiments were previously presented at the 2006 AIAA/MST, 2009 ICNS, and 2011 ISAP conferences.

References

[1] SESAR Consortium, SESAR Master Plan. April 2008. Available at: www.atmmasterplan.eu.

[2] EMMA2 Consortium, "A-SMGCS services, procedures, and operational requirements (SPOR)," IP of 6th FP of European Commission/DG TREN, Brussels, Belgium, 2008. Available at: www.dlr.de/emma2

[3] Joint Planning and Development Office, "Concept of Operations for the Next Generation Air Transport System, v2.0," June 13, 2007. http://www.jpdo.gov/library/NextGen_v2.0.pdf

[4] Y. C. Jung, T. Hoang, J. Montoya, G. Gupta, W. Malik, and L. Tobias, "A concept and implementation of optimized operations of airport surface traffic," 10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Fort Worth, TX, September 2010.

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

[6] S. Rathinam, J. Montoya, and Y. Jung, "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, September 2008.

[7] V. H. L. Cheng, V. Sharma, and D. C. Foyle, "A study of aircraft taxi performance for enhancing airport surface traffic control," IEEE Transactions on Intelligent Transportation Systems, Vol. 2, pp. 39-54, 2001.

[8] D. C. Foyle, B. L. Hooey, C. L. Kunkle, M. F. J. Schwirzke, and D. L. Bakowski, "Piloted simulation of NextGen time-based taxi clearances and tailored departures," Proceedings of the 2009 IEEE/AIAA Integrated Communications, Navigation and Surveillance Conference (ICNS). Arlington, VA, May 2009.

[9] B. L. Hooey, D. C. Foyle, and A. D. Andre, "Integration of cockpit displays for surface operations: The final stage of a human-centered design approach," SAE Transactions: Journal of Aerospace, Vol. 109, pp. 1053-1065, 2000.