The International Journal of Aviation Psychology

ISSN: 1050-8414 (Print) 1532-7108 (Online) Journal homepage: http://www.tandfonline.com/loi/hiap20

Flight-Deck Surface Trajectory-Based Operations

David C. Foyle, Becky L. Hooey, Deborah L. Bakowski & Christina L. Kunkle

To cite this article: David C. Foyle, Becky L. Hooey, Deborah L. Bakowski & Christina L. Kunkle (2015) Flight-Deck Surface Trajectory-Based Operations, The International Journal of Aviation Psychology, 25:2, 77-96, DOI: 10.1080/10508414.2015.1097090

To link to this article: http://dx.doi.org/10.1080/10508414.2015.1097090

Published online: 02 Feb 2016.

THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, *25*(2), 77–96
ISSN: 1050-8414 print / 1532-7108 online
DOI: 10.1080/10508414.2015.1097090

Routledge
Taylor & Francis Group

Flight-Deck Surface Trajectory-Based Operations

David C. Foyle,[1] Becky L. Hooey,[2] Deborah L. Bakowski,[2] and Christina L. Kunkle[2]
[1]*Human Systems Integration Division, NASA Ames Research Center*
[2]*San Jose State University at NASA Ames Research Center*

The results of three piloted simulations investigating flight-deck surface trajectory-based operations (STBO) are presented. Commercial transport pilots were given taxi clearances with time and speed components on the primary flight display and were required to taxi to the departing runway or intermediate intersections. Results show that when pilots were provided with speed-only taxi clearances, pilots either had poor required time of arrival (RTA) conformance with acceptable estimates of attentional distribution and safety, or had good RTA conformance with unacceptable attentional distribution and safety estimates. A flight-deck error-nulling algorithm/display allowed pilots to conform accurately with taxi RTA clearances while maintaining safety. Results are discussed in terms of pilot multitasking in the busy airport surface operations environment.

On a global basis, research is underway to design the next generation of airspace systems to increase capacity and throughput in all weather conditions, and reduce emissions and pollution, by taking advantage of new technology and concepts in both air traffic management and the flight deck. The SESAR (SESAR Consortium, 2008) and EMMA2 (EMMA2 Consortium, 2008) efforts in Europe and the NextGen (Joint Planning and Development Office, 2010; National Academies Press, 2015) efforts in the United States are core programs of these new technology efforts. Under these programs, all phases of flight are being investigated: preflight, push back, taxi, takeoff, departure, climb, en-route cruise, descent, approach, landing, taxi, and gate arrival. The studies reported here investigated the taxi-out departure environment (from the ramp area to the runway) in the NextGen air transportation system environment.

NEXTGEN SURFACE TRAFFIC MANAGEMENT SYSTEMS

Current National Aeronautics and Space Administration (NASA) research efforts are aimed at the development of surface traffic management (STM) systems for air traffic control (ATC) to provide optimized taxi clearances enabling efficient airport traffic operations and improving throughput. In addition to having specified taxi arrival times at

Correspondence should be sent to David C. Foyle, MS 262-4, NASA Ames Research Center, Moffett Field, CA 94035. E-mail: David.C.Foyle@nasa.gov

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 (see Cheng, 2002). 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 pilots to be active participants in reaching specified locations at specified times. Future taxi clearances would have a speed- or time-based component with which the pilot must comply —these NextGen taxi operations have been referred to as "4D taxi" (with the fourth dimension referring to the time component), or surface trajectory-based operations (STBO). In contrast to current-day taxi operations, which do not require pilots to taxi at any specified speed or with specified time requirements, future STBO 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 (Cheng, Yeh, Diaz, & Foyle, 2004; Rathinam, Montoya, & Jung, 2008). Cheng, Sharma, and Foyle (2001) showed 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 (Cheng et al., 2001).

There are many variants of the STBO concept. In the most highly constrained STBO case, the aircraft is required to follow a fully defined time and location profile, such that every (x, y) point along the taxi clearance route has a defined predictable time (t) associated with it (i.e., x_t, y_t). A simpler, less constrained STBO case is when only one or two airport surface locations along the taxi clearance route have required times of arrival (RTAs). For example, arrival at the departure runway might be the only point in the taxi clearance that has a defined arrival time. 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 intersections, active runway crossings, etc.) at specific times (i.e., time constraint points). The aircraft's commanded speed or time might need to be adjusted if the pilot is unable to conform to the STBO command, 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.). Because the STBO taxi concept is in its infancy, current efforts aim to affect the design of the underlying ATC STM algorithms, so that the resulting STM system does not exceed pilot and aircraft performance capabilities. Currently there is no accepted STBO RTA accuracy design 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 (Cheng et al., 2004). In the future NextGen environment, 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.

STBO IMPLICATIONS FOR PILOT PERFORMANCE

As a minimum requirement, future implementations of STBO taxi clearances will require new information on the flight deck (i.e., coordinated time information, airport taxiway distances, etc.) and likely will require advanced displays to support pilots during taxi operations. Care must be taken in the development of these new flight-deck systems to ensure that they are effectively integrated into existing pilot tasks. Contrary to what one might expect, taxi navigation itself is a demanding operation. Interviews with individual pilots and focus groups have indicated that current taxi-out and departure are the busiest phases of flight for the flight crew. During taxi-out, crew taxi operations include maneuvering the aircraft, maintaining separation from other aircraft and vehicles, navigating the taxi clearance by referring to airport signage and the airport taxi chart, and communicating with ATC regarding the clearance. Billings (1997) characterized these multiple tasks for the pilot as those of aviation, navigation, communication, and the management of resources. Theunissen, Rademaker, Jinkins, and Uijt de Haag (2002) described a detailed analysis of the navigation task and information requirements when pilots are conducting taxi operations with an airport moving map (AMM). They described the cognitive and task complexities associated with the taxi navigation task, detailing the processes of information collection, integration, and extrapolation required of pilots to safely and efficiently navigate on the airport surface. Relatedly, Hooey and Foyle (2006) analyzed taxi navigation errors using a taxonomy of errors that included errors related to the underlying processes of navigational planning, decision making, and navigational execution. In addition to these taxiing duties, the flight crew also conducts 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 counts and baggage weight loads, communicating with cabin crew and passengers, and completing predeparture briefings related to the normal departure and safety backup procedures in the case of such off-nominal events as an engine-out on takeoff.

It is clear from the preceding description that when taxiing an aircraft, pilots are working in a multitasking environment, monitoring systems, making control inputs, and planning. The multitasking environment, where users are required to manage a variety of tasks while time-sharing cognitive resources among them, has been a topic of research over the last few decades. Early work by Norman and Bobrow (1975) proposed that, in the multitasking environment, performance limits can occur because of either data quality limitations or limits in the amount of cognitive processing resources available. Wickens (1984; also see Wickens, 2002) proposed multiple resource theory, where rather than having a single cognitive resource pool that is available for cognitive processing, users have a collection of resource pools that can be applied to processes serially or in parallel, depending on the task demands and requirements. Increases in workload and associated performance decrements are found when multiple task demands require a single cognitive resource simultaneously. In addition to data limitations and cognitive resource limitations that affect performance, similarly strategic factors affect how people interleave their attention when multitasking. Janssen and Brumby (2010), in a driving task where drivers also had to concurrently dial a phone, found that rather than switching between driving and dialing at the natural telephone number formatting breaks (i.e., city code, prefix, etc.), drivers switched attention between the two tasks in such a manner as to maintain performance integrity on the higher priority task (i.e., acceptable lane control) while dialing. The authors took this as evidence of the importance of strategic factors (rather than task structure factors) that allow users to effectively allocate their attention in multitasking environments to maintain performance on high-priority tasks (i.e., driving in this case).

Under the STBO concept, pilots would be required to interleave the new time requirement (i.e., RTA) task into the already demanding list of current taxi navigation tasks. The aircraft maneuvering, separation, and navigation tasks primarily require the captain (typically responsible for aircraft taxi) to be “eyes-out”; that is, viewing the airport environment through the flight deck windows. Adding an STBO RTA task in which speed or time must be adhered to would require the captain to add an additional “eyes-in” task in which the speed or time information must be monitored via the flight deck’s avionics, thus requiring the captain to switch attention between the eyes-out tasks and the eyes-in tasks.

As a first step toward defining the information requirements for flight deck systems to support this new task, Williams, Hooey, and Foyle (2006) conducted a study to determine the relative contribution of speed and time information on pilot taxi performance and how pilots use that information to meet taxi RTAs. In a medium-fidelity flight-deck simulator, 18 commercial airline pilots were required to complete taxi routes while achieving specific average speeds or completion times when speed, time, or both speed and time information were available on a simulated head-up display (HUD). For the speed format, pilots were instructed to taxi to a runway crossing at a commanded average speed using a digital readout of the current ground speed and commanded speed. With the time format, pilots were instructed to taxi to a specific runway crossing arriving at a commanded RTA using a digital readout of the elapsed time and the RTA. In the speed and time format, both speed and time information were available and pilots were required to comply with both a required speed and its corresponding commanded RTA. The commanded speeds and times were based on combinations of four required speeds (10, 14, 18, and 22 kt) and three route distances (3,000, 6,000, and 12,000 ft). The results showed a significant interaction between route distance and required speed, with larger RTA error as the route distance increased from 3,000 ft (M = .002 sec early) to 6,000 ft (M = 4.87 sec early) to 12,000 ft (M = 11.34 sec early) for commanded taxi speeds of 10, 14, and 18 kt but not for the fastest commanded taxi speed of 22 kt. The results also revealed larger RTA error when required speeds were either slower or faster than is typical during normal taxi operations (i.e., 10 kt, M = 22.59 sec early; 22 kt, M = 9.34 sec late).

This research has two goals: to determine (a) if RTA conformance can be improved by manipulating information format and flight-deck procedures for controlling the speed of the aircraft to meet the required RTA; and (b) how to best support pilots’ interleaving of the new time-based taxi task with existing taxi task demands. To these ends, three experiments investigating RTA conformance were conducted. Experiment 1 evaluated pilots’ ability to meet a taxi RTA by following the simplest procedural solution, a verbal speed command from ATC without any conformance requirements. Experiment 2 also evaluated a procedural solution with ATC supplying a verbal speed requirement, but with an added aircraft conformance requirement for speed and acceleration or deceleration. Finally, Experiment 3 tested an error-nulling algorithm/display that informs the pilot of the required speed, dynamically adjusting based on pilot performance for the duration of the route to meet the RTA. The results of the three experiments are discussed in the context of the multitasking environment that pilots face during taxi operations; that of maintaining aircraft separation, controlling and navigating the aircraft, as well as the new task of meeting the taxi clearance RTA. Objective and subjective measures of taxi performance, attention allocation, and safety are analyzed and discussed within the context of the three experiments as the task demands of RTA conformance increase from simple procedures in Experiment 1, to more complex procedures in Experiment 2, ending with an underlying algorithm and display solution that was developed to support strategic usage in Experiment 3.

EXPERIMENT 1: COMMANDED SPEED—WITH INTERMEDIATE TIME CONSTRAINT POINTS

Experimental Objective

The simplest STBO 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 as part of the taxi clearance. This experiment replicates the Williams et al. (2006) study described previously with the addition of intermediate time constraint points (and new speed requirement) and higher fidelity flight-deck displays. In this experiment, the ground controller provided the pilot a taxi clearance with a required speed to maintain. Pilots in this experiment were not required to follow any specific acceleration or deceleration speed profiles. Given the distance effect (RTA error increased as route length increased) observed by Williams et al. (2006), it was predicted that intermediate time constraint points would allow aircraft to reach the (shorter) separate RTAs more accurately, resulting in overall better RTA performance. When following a commanded speed, the pilot only knows the current deviation from that speed, and does not know the cumulative effect of speed error. Adding intermediate time constraint points to a longer taxi route breaks up a long route into shorter segments each with its new RTA. Reaching the time constraint point resets the previous time error to zero—thus, one would expect that adding time constraint points would lead to reduced RTA error at those points.

Method

Participants. Eight commercial pilots (six captains, two first officers), current or recently retired, participated in the study. Mean pilot age was 49 years, and mean flight hours logged was 5,029 hr. One pilot was female and seven were male.

Flight simulation. The study was conducted in a medium-fidelity simulator with Boeing 737 modeled dynamics in the Human-Centered Systems Laboratory (HCSL) at the NASA Ames Research Center (the same simulator used by Williams et al., 2006). The simulated environment modeled Dallas/Fort Worth (DFW) International Airport with 1,200 ft visibility. The forward out-the-window scene was rear projected with visual angles: 53.1° (Horizontal, H) by 41.1° (Vertical, V). Side window views were displayed on two side monitors 31.9° (H) by 24.2° (V). Pilots controlled the simulated aircraft using a tiller, throttle, and rudder with toe brakes. The simulator flight deck included a primary flight display (PFD), navigation display (ND), AMM, datalink, and electronic checklist. The PFD was modified for taxi operations—the speed scale was active and rescaled to 0 to 60 kt to support taxi operations. The commanded ground speed was displayed both digitally in magenta directly above the speed tape and as a magenta analog pointer (“speed bug”). Current speed was shown as a sliding indicator with the digital value inside. To assist in airport navigation, the simulator flight deck included a dynamic AMM that

FIGURE 1 Airport moving map (AMM) showing route and ownship position (chevron).

depicted the airport layout. At the start of each trial the AMM showed the entire airport in north-up view to support route planning. It changed to track-up perspective mode (Figure 1) 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 AMM, indicated the cleared route with positive cleared-to-cross runway clearance.

Experimental design. The experiment was a within-participant design with two factors, number of time constraint points (1, 3, or 5) and commanded speed (10, 14, 18, or 22 kt). 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 time constraint point values × 4 speeds × 2 repetitions).

Procedure. Pilots completed departure taxi scenarios from a ramp departure spot to a departure runway (mean taxi time = 9.5 min). At the beginning of each trial, the experimenter, acting as ground controller, issued a verbal taxi clearance to the departing runway that also appeared in text and graphically on the AMM. After the pilot reviewed the map and clearance, the AMM switched to the track-up perspective view (see Figure 1), and the trial began. Data were collected after the pilot had completed four 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. With each new commanded speed, the digital and the graphic speed indicator showing commanded speed on the PFD changed to the new commanded speed value, and the pilot was required to change speed accordingly. While taxiing, pilots received a data-linked departure clearance for verification, and were instructed to monitor the status of an electronic checklist to emulate crew tasks. Pilots were instructed to prioritize the tasks as follows, from highest to lowest: aircraft separation, taxi clearance route navigation, taxi speed, electronic checklist monitoring, and departure clearance verification.

Results and Discussion

The primary measure of pilot performance on the taxi task was RTA error, calculated by subtracting the RTA from the observed arrival time at each traffic constraint point. 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. Pilots did not receive an explicit commanded RTA—instead, they received a commanded speed that they were required to follow on straight segments, told not to exceed 15 kt in turns, and instructed to accelerate or 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 kt/sec acceleration and deceleration; assumed turn speeds were 15 kt for commanded speeds of 18 and 22 kt, and equal to the straight segment speed for commanded speeds of 10 and 14 kt.

A 3 (number of time constraint points) by 4 (commanded speed) within-participant analysis of variance (ANOVA; see Figure 2) revealed a significant main effect of commanded speed, $F(3, 21) = 24.87$, $p < .001$. Paired t-tests were conducted using a Bonferroni-adjusted alpha of .008 (.05/6). There was more positive RTA error, indicating pilots arrived at the queue later, when commanded taxi speed was 22 kt or 18 kt than when the commanded speed was 14 kt, $t(7) = 8.97$, $p < .001$; $t(7) = 5.29$, $p = .001$, respectively; and compared to when the commanded speed was 10 kt, $t(7) = 5.78$, $p = .001$, $t(7) = 4.32$, $p = .004$, respectively. No other pairwise comparisons were significant.

These results must be interpreted within the context of a significant interaction between the number of time constraint points and commanded speed, $F(6, 42) = 6.79$, $p < .001$. Simple main effect tests were conducted to evaluate the effect of number of time constraint points at each level of commanded speed. Significant simple main effect tests were followed up with paired t-tests using a Bonferroni-adjusted alpha of .017 (.05/3). When the commanded speed was 22 kt, RTA error differed as a function of number of time constraint points, $F(2, 14) = 39.62$, $p < .001$. Specifically, at 22 kt commanded speed, RTA error was highest when there was only one time constraint point and error was reduced when there were three and five time constraint points, $t(7) = 6.22$, $p < .001$, $t(7) = 6.46$, $p < .001$, respectively; RTA error observed with five time constraint points was reduced compared to that with three time constraint points, $t(7) = 4.36$, $p = .003$. RTA error did not differ

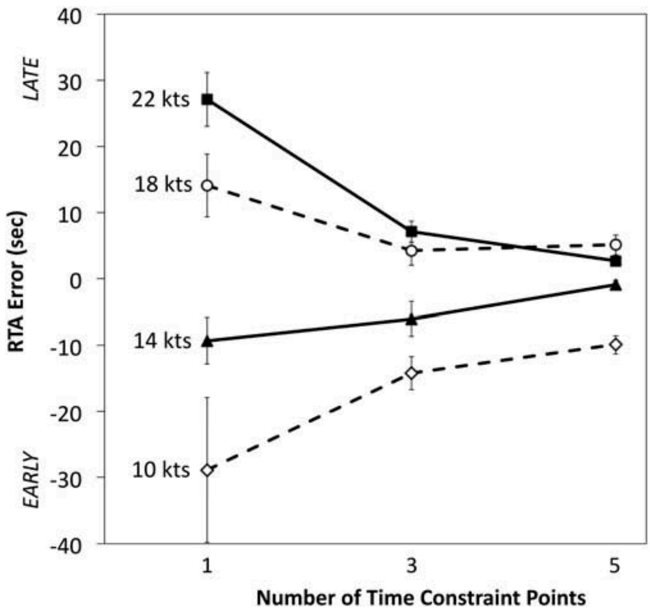

FIGURE 2 Experiment 1: Mean required time of arrival (RTA) error as a function of commanded speed and number of time constraint points. Error bars = ±1 *SE*.

significantly as a function of the number of time constraint points when commanded speeds were 10, 14, or 18 kt using the Bonferroni correction, although the simple main effect at commanded speed of 18 kt was significant, $F(2, 14) = 6.34$, $p = .011$.

In this experiment, pilots were to follow the taxi route clearance with the associated ATC-commanded speed. With speed commands only, pilots exhibited more difficulty maintaining a relatively fast taxi speed (22 kt) for a long distance, as in the one time-constraint point condition—but RTA error was reduced for these speeds by adding three or five time constraint points that served to decrease the distance of each segment. This route-distance effect was similar to that seen in Williams et al. (2006). RTA error was closest to zero when taxi speed was 14 kt, 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 kt, because they did not need to slow down for turns and then resultantly correct for the slower speed.

EXPERIMENT 2: COMMANDED SPEED WITH SPEED AND ACCELERATION PROFILES

Experimental Objective

Following up on the previous results, Experiment 2 was aimed at determining if RTA error would be reduced if pilots were explicitly required to follow the speed acceleration and deceleration profile used to calculate the route RTA. In STBO operations, this would make the ownship's speed, and hence location, performance more predictable. Pilots in Experiment 2 were again required to follow commanded straight and turn speeds, but in contrast to the previous experiment, pilots in this experiment were required to follow specific acceleration and deceleration speed profiles. It was predicted that requiring pilots to follow specific acceleration and deceleration speed profiles would lead to reduced RTA error as compared to Experiment 1. In addition, the impact of a target speed deviation limit around the commanded speed was assessed. Inasmuch as larger pilot speed control deviations are reflected in increased RTA error, it was expected that such a target speed deviation limit around the commanded speed would lead to relatively low RTA error because input speed would more closely match the required speed.

Method

Participants. Eighteen commercial pilots (13 captains, five first officers), current or recently retired, participated in the study. Mean pilot age was 45 years and mean flight hours logged was 3,832 hr. One pilot was female and 17 were male.

Flight simulation. The study used the same simulator and visibility conditions (high visibility and distant fog or haze) at DFW airport as in Experiment 1. For this experiment, the two left and right side monitors subtended 29.6° visual angle. The same PFD (except that no commanded speed was displayed; commanded speeds were provided verbally by ATC) and AMM configuration as in Experiment 1 were used. While taxiing, pilots wore an Applied Science Laboratory Model 501 head-mounted eye tracker.

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 kt). 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 kt 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 actual operations.

Procedure. Pilots completed departure taxi scenarios from a ramp departure spot to a departure runway. A computer-triggered prerecorded verbal ATC command providing a changed taxi speed (e.g., “NASA 227, taxi at 14 kt”) accompanied each taxi segment transition (14, 18, or 22 kt) at the time consstraint point. Segment distances and speed changes were not depicted on the AMM. In addition to the speed command that pilots received at the start of each segment, two specific aircraft control and 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 and speed profile instructions were as follows:

1. Taxi all turns at 14 kt.
2. Accelerate or decelerate at 2 kt/sec (e.g., 7 sec acceleration for 0 kt to 14 kt; 2 sec for a speed change from 22 kt to 18 kt, etc.).

The final segment in each trial ended several hundred feet before the threshold of the departure runway. Average trial length was 8 min, 42 sec.

Results and Discussion

RTA error. The 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. The primary measure of pilot performance on the taxi task was RTA error, calculated by subtracting the RTA from the observed arrival time at each traffic constraint point. The RTA values were calculated using the nominal or instructed speed profile, considering the number of turns, nominal turn speed of 14 kt, and 2 kt/sec acceleration or deceleration before and after turns, for initial taxiing, and for commanded speed changes.

As seen in Figure 3, RTA error is quite low in all conditions, and is much less than that found in Experiment 1 (note reduced scale compared to Figure 2) where pilots were not required to

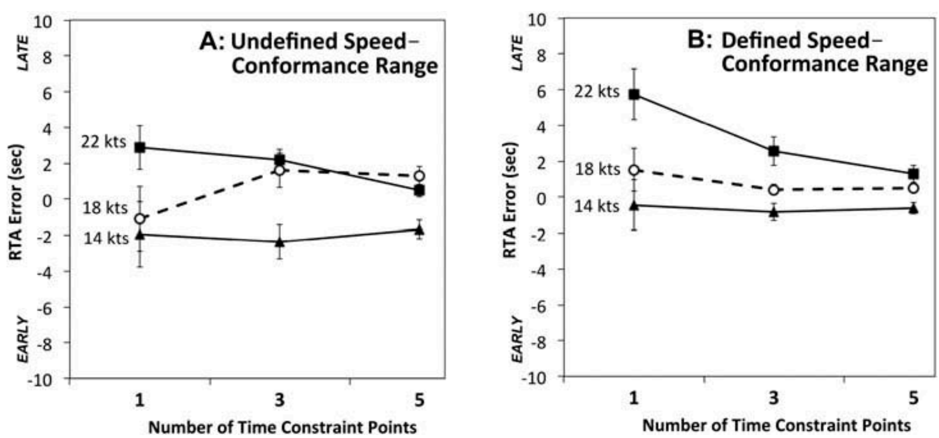

FIGURE 3 Experiment 2: Mean required time of arrival (RTA) error as a function of commanded speed and number of time constraint points for the (a) undefined and (b) defined speed-conformance conditions. Error bars = ±1 *SE*.

follow specific speed and acceleration or deceleration profiles. A 2 (speed-conformance implementation) by 3 (number of time constraint points) by 3 (commanded speed) within-participant ANOVA was conducted. Although the main effect of speed-conformance implementation was not significant, there was an interaction between speed-conformance implementation and number of time constraint points, $F(2, 34) = 4.44$, $p = .019$. Post-hoc tests showed a simple main effect of number of time constraint points in the defined speed-conformance condition, $F(2, 34) = 7.75$, $p = .002$. RTA error for one-segment trials was significantly higher than for three-segment trials, $t(17) = 2.70$, $p = .015$; and five-segment trials, $t(17) = 3.21$, $p = .005$. This was consistent with Experiment 1, in which pilots exhibited more difficulty maintaining a commanded taxi speed for a long distance (as in the one time-constraint-point trials) than for shorter distances (three- or five-time-constraint-point trials). The simple main effect of number of time constraint points in the undefined condition was not significant.

There was also an interaction between number of time constraint points and commanded speed, $F(4, 68) = 3.44$, $p = .013$. Simple effect tests revealed a significant difference in RTA error as a function of the number of time constraint points when the commanded speed was 22 kt, $F(2, 34) = 8.45$, $p = .001$: Paired *t*-tests with Bonferroni-adjusted alpha of .017 (.05/3) revealed that RTA error was significantly lower when there were five time constraint points than when there were three time constraint points, $t(17) = 3.53$, $p = .003$, and one time constraint point, $t(17) = 3.61$, $p = .002$; RTA error for trials with only one time constraint point did not differ significantly from trials with three time constraint points. RTA error did not differ significantly as a function of the number of time constraint points when the commanded speed was 14 or 18 kt.

Percent dwell time on PFD speed. The percentage of time that the pilots fixated (percent dwell time) on the current-speed read-out displayed on the PFD was recorded. 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, $F(2, 20) = 41.29$, $p < .001$. Compared to the current-day baseline condition of 7.55 ($SE = 1.03$) mean percent dwell time, for the two speed-conformance conditions, pilots spent 17.94 ($SE = 2.06$) mean percent dwell time in the undefined speed-conformance condition and 24.39 ($SE = 2.65$) mean percent dwell time in the defined speed-conformance condition of the trial looking at the speed display. All pairwise comparisons were significant: undefined versus defined, $t(10) = 4.95$, $p < .05$; undefined versus baseline, $t(10) = 5.82$, $p < .05$; defined versus baseline, $t(10) = 7.04$, $p < .05$. In absolute as well as relative terms, these are large percentages of the trial to be looking 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 taxiing aircraft. In addition, pilots were asked, “Would the demand of having to maintain the required speed conformance range compromise safety in the real world?” in a questionnaire. A chi-square goodness-of-fit test revealed that significantly more pilots ($n = 14$) responded that the demand of having to maintain the required speed conformance range in the real world would compromise safety than responded that it would not ($n = 4$), $\chi^2(1, N = 18) = 5.56$, $p < .05$.

Data from this experiment indicated that when pilots are given speed-based taxi clearances and are required to control their aircraft according to precise acceleration or deceleration speed profiles, RTA conformance is quite good (on average 6 sec or better). However, associated with this good RTA conformance, pilots spent an inordinate amount of time viewing and tracking their speed and felt that it could be unsafe. Given that taxiing relies so heavily on out-the-window airport navigation and aircraft, vehicle, and pedestrian separation, the eye-tracking data raise safety concerns that are addressed later in this article.

EXPERIMENT 3: ERROR-NULLING ALGORITHM/DISPLAY

Experimental Objective

The previous two experiments assessed RTA error when pilots were given taxi clearances with speed requirements and a flight deck equipped with simple readout displays of current speeds. Experiments 1 and 2 demonstrated that, when requiring the pilots to follow a commanded speed, adding time constraint points to the routes decreased RTA error, but at a cost of increased eyes-in time and estimated safety. One of the contributing factors to RTA error with commanded speed only is that pilots do not know the amount of accumulating speed error as they progress along the taxi route. To counter this, Experiment 3 was aimed at developing a flight-deck STBO display that offloaded the nulling of accumulating RTA error to the avionics. In so doing, this would allow the pilot to attend visually to the display when task demands permit; that is, it would support multi-tasking. Such a design solution enables pilots to meet the taxi route RTA without moment-by-moment tracking of ground speed. As mentioned previously, the tasks of maneuvering the aircraft, navigating, and maintaining separation are largely eyes-out (looking out the window) taxi tasks, whereas speed maintenance is primarily an eyes-in task, requiring the monitoring of avionics and subsequent control of speed. Furthermore, under current-day operations only rarely is there a need to consult a speed readout while taxiing (and only the most modern aircraft have ground speed indicators suitable for taxi). Specifically, the 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 or her attention appropriately and to switch tasks strategically among aircraft separation, airport navigation, and the many other concurrently required flight-deck tasks. Clearly, a poorly designed display requiring large amounts of visual scrutiny, mental calculations, or cognitive interpretation would not enable the pilot to interleave all of the tasks required in an efficient manner.

In this study, pilots were provided a PFD speed display driven by an error-nulling algorithm that computed the necessary speed to arrive at the required location (intersection or runway) at the RTA. It was predicted that the error-nulling algorithm display would lead to relatively low RTA error for all speeds and number of time constraint points.

Method

Participants. Eight male commercial pilots (seven captains, one first officer), current or recently retired, participated in the study. Mean pilot age was 42 years; mean flight hours logged was 6,143 hr.

Flight simulation. The same physical setup, displays (with the exception of the PFD), and general procedure were used as reported in Experiment 1. The PFD was nearly identical to that used in Experiment 1, with the exception that time information was included (see Figure 4) and

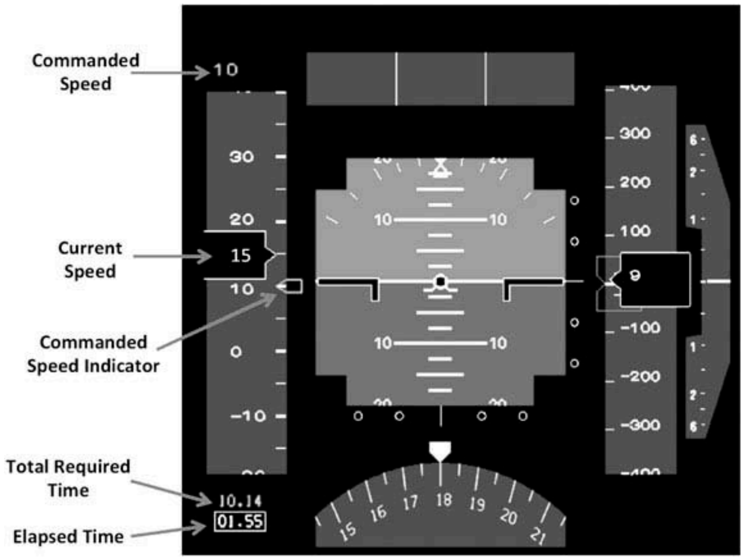

FIGURE 4 Error-nulling algorithm/display. Speed tape (left side) shows commanded speed = 10 kt, current speed = 15 kt, elapsed taxi time = 1:55 min/sec, total required taxi time = 10:14 min/sec. Commanded speed is dynamically calculated and updated by error-nulling algorithm.

the commanded speed adjusted dynamically. The lower left area of the PFD included elapsed time in a white box (in min, sec), counting upward from zero, and RTA in magenta (in min, sec). An error-nulling algorithm dynamically compensated for speed maintenance errors by continually adjusting the commanded speed indicator based on the remaining RTA and remaining distance to the time constraint point according to the calculation:

$$\text{Commanded Speed} = \text{Remaining Distance}/\text{Remaining Time}$$

Thus, by following the currently indicated 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 the RTA error).

The AMM was identical to that used in Experiment 1, with the exception that the time constraint point location was shown graphically on the AMM as a yellow bar across the cleared taxi route.

Experimental design. The same factors and routes were used as in Experiment 1. The experiment was a within-participant design with two factors: number of time constraint points (1, 3, or 5) and commanded speed (10, 14, 18, or 22 kt). Similarly, there were a total of 24 taxi trials (3 time constraint point values × 4 commanded speeds × 2 repetitions) using the same 12 unique taxi routes.

Procedure. Pilots completed departure taxi scenarios from a ramp departure spot to a departure runway following the same general procedures as in Experiment 1. However, when reaching a time constraint point, the pilots received an auditory chime with the verbal cue, "Checkpoint."

Results and Discussion

The primary measure of pilot performance on the taxi task was RTA error, calculated by subtracting the RTA from the observed arrival time at each traffic constraint point. It should be noted that mean RTA error was negligible, less than 10 sec in all conditions, well within the ±15 sec window that might be required in future STBO operations as discussed earlier. A 3 (number of time constraint points) by 4 (commanded speed) within-participant ANOVA (see Figure 5) revealed an interaction between number of time constraint points and commanded speed, $F(6, 42) = 3.67$, $p = .005$. Simple main effect tests were conducted to evaluate the effect of number of time constraint points at each level of commanded speed and were followed-up with paired *t*-tests. The effect of number of time constraint points was significant only when the commanded speed was 22 kt, $F(2, 14) = 5.03$, $p = .023$; however, pairwise *t*-test comparisons failed to reach significance using the Bonferroni-adjusted alpha of .017 (.05/3).

As can be seen in Figure 5, the availability of an error-nulling algorithm driving a commanded-speed display allowed pilots, on average, to reach the runway or time constraint point within approximately 10 sec of the RTA for all commanded speeds and number of time constraint points tested (similar to the results of Experiment 2). Good end-route RTA

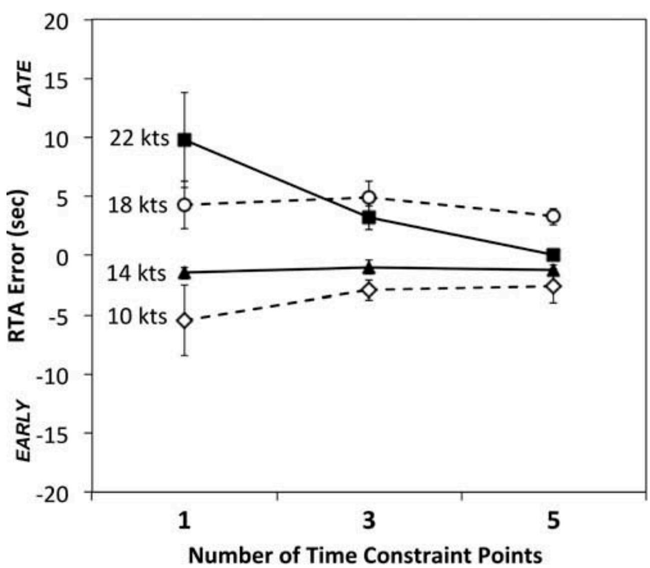

FIGURE 5 Experiment 3: Mean required time of arrival (RTA) error as a function of commanded speed and number of time constraint points. Error bars = ±1 *SE*.

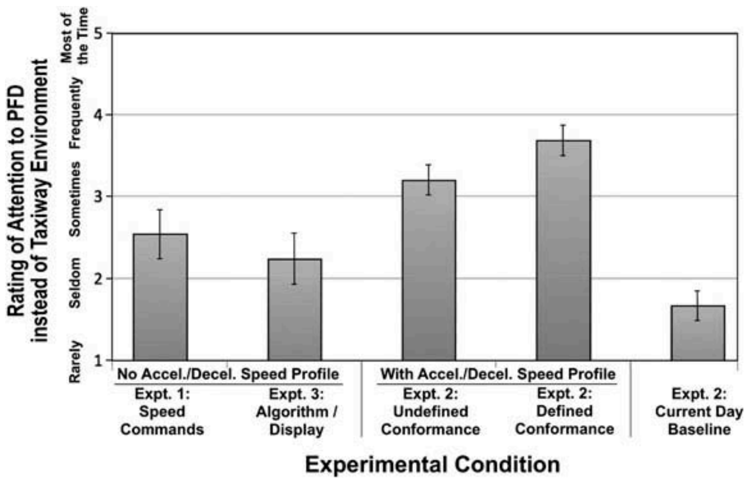

FIGURE 6 Mean rating response to posttrial question, "How often did you find yourself focusing on the PFD [primary flight display] speed or time display, when you should have been paying attention to the external taxiway environment?" Error bars = ±1 *SE*.

performance was also found in an STBO experiment (Jones, Prinzel, Bailey, Arthur, & Barnes, 2014) that used error-nulling advised speed, but presented different flight-deck displays than tested here. As discussed in the next section (and seen in Figure 6), along with the good RTA performance observed in Experiment 3, pilots reported that the use of the algorithm/display did not interfere with their attention to the airport environment. In Experiment 3, the nature of the algorithm is such that moment-by-moment attention to speed is not required, and thus the pilot can use it strategically, attending to the current commanded speed when workload permits. Momentary speed fluctuations have relatively little effect on the dynamically calculated speed because it is averaged over the remaining portion of the taxi route.

USAGE AND SAFETY IMPLICATIONS

The accurate taxi RTA data presented (especially in Experiments 2 and 3) is not the whole story regarding flight-deck STBO. New flight-deck requirements can put new pressures on the crew, especially if not designed with consideration of the multitasking context. It was observed that pilots in all three experiments were able to maintain performance in the two highest priority tasks—the taxi navigation task (following the taxi route) and maintaining aircraft separation—while performing the third-priority RTA speed task. However, the implementation method of the speed task determined how well they performed on that task and how much of their visual attention the task demanded. Specifically, the error-nulling algorithm/display in Experiment 3 allowed pilots to effectively interleave the speed maintenance task with the other taxi tasks without sacrificing performance on the speed task or safety. However, the other implementation approaches that required pilots to continuously track commanded speed did not support effective task interleaving and resulted in either reduced performance on the speed task (Experiment 1) or reduced safety as operationalized by excessive visual attention demands (Experiment 2; i.e., pilots spent a large percentage, 18–24%, of time eyes-in, viewing their ground speed, compared to what they would do under current-day operations, 8%).

Similarly, posttrial and poststudy questions also inform potential safety issues. In Experiment 2, pilots were asked the following poststudy question: "Would the demand of having to maintain the required speed-conformance range compromise safety in the real world?" Significantly, the data showed that 14 of the 18 pilots tested responded that the demand of having to maintain the required speed-conformance range would compromise safety. Taken together, the eye-tracking data and the safety question raise safety concerns for conditions when pilots must precisely follow speed profiles.

In all three experiments, pilots were asked to rate on a scale from 1 (*rarely*) to 5 (*most of the time*; see Figure 6) their response to the posttrial 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 Figure 6, pilots reported looking at the displays instead of paying attention to the external taxiway environment more often in Experiment 2 (which required pilots to taxi according to a specific acceleration and deceleration profile) than in either Experiment 1 or Experiment 3, both of which did not require following the acceleration or deceleration speed profiles: Experiment 1 versus Experiment 2 defined conformance, $t(24) = 3.359$, $p < .01$; Experiment 1 versus Experiment 2 undefined conformance, $t(24) = 1.95$, $p < .05$, 1-tail significance; and Experiment 3 versus Experiment 2 defined conformance, $t(24) = 4.169$, $p < .001$; and Experiment 3 versus Experiment 2 undefined conformance, $t(24) = 2.78$, $p < .05$; all statistics are significant 2-tailed except as noted. These data are further evidence of the challenging nature of STBO taxiing when pilots are required to follow a taxi speed profile with specified acceleration or deceleration precisely.

The posttrial attention question ratings are shown in Figure 7 as a function of the corresponding mean RTA errors for the speeds tested across the three experiments (one 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 3 (with the error-nulling algorithm/display) lie within the "optimal zone," suggesting that only the error-nulling algorithm/display 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 might affect safety precursors (e.g., time looking out the window). Clearly, the present data suggest that STBO flight-deck procedures that require precise tracking of acceleration and deceleration speed profiles yield increased safety concerns, whether measured with eye-tracking data and rating assessment.

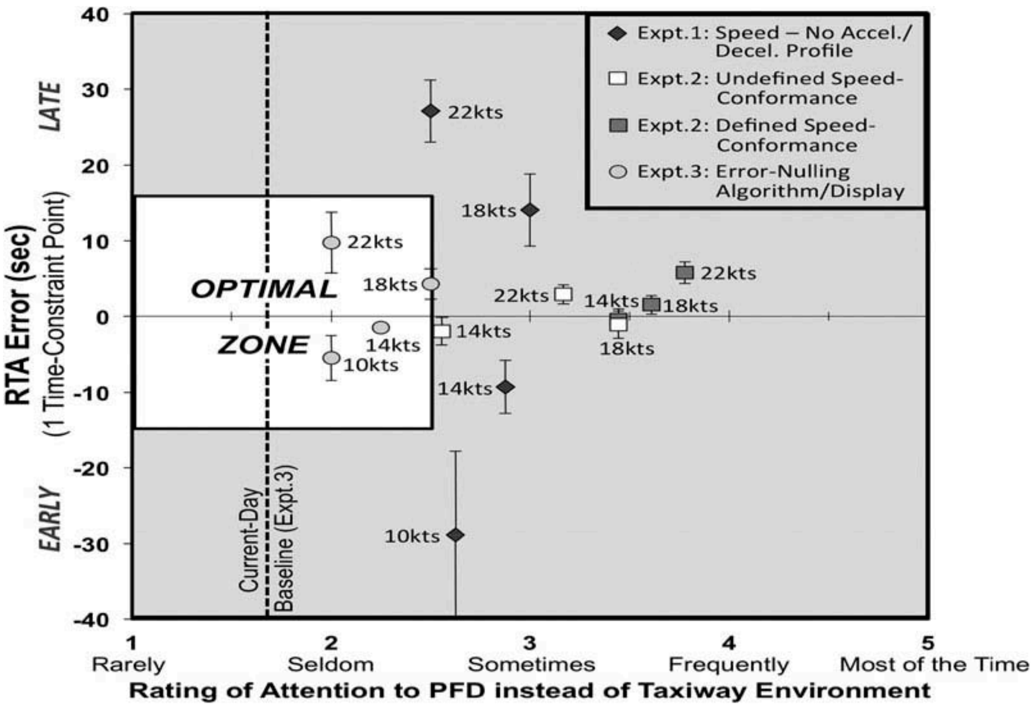

FIGURE 7 Mean ratings for the attentional allocation question (from Figure 6) by mean required time of arrival (RTA) errors (one time-constraint point only) for the commanded speeds tested in Experiments 1, 2, and 3. (Dashed line represents "current-day baseline" condition rating.) Error bars = ±1 *SE*.

GENERAL DISCUSSION AND SUMMARY

A series of three 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. Experiment 1 was an extended replication of Williams et al. (2006), with the addition that the taxi clearances had one, three, or five time constraint points, each with its own speed and time requirement. Both studies found evidence of a route-distance effect in which RTA error was greatest for longer distances and reduced when the route was divided into shorter segments. In Experiment 2, RTA conformance was quite good (better than 6 sec on average) when pilots were required to follow a precise speed and acceleration and deceleration profile. However, this demanded higher levels of visual attention (18–24% dwell time). Pilots rated that more eyes-out (looking out the window) time was required, and, more pilots (14 out of 18) rated the display "unsafe." Experiment 3 showed that the error-nulling algorithm/display can mitigate these concerns, and provide good RTA performance without the attentional and potential safety costs seen in Experiment 2.

Flight Deck Implications

Taxi is a busy phase of flight and therefore adding any new duties and tasks such as RTA conformance must be done carefully, taking pilot workload, task allocation, and safety into account. New flight-deck displays to support the added task of an RTA requirement must be designed carefully to support efficient pilot multitasking. As discussed earlier, the authors consider the error-nulling algorithm/display tested in Experiment 3 to enable multitasking because it can be used strategically by the pilot without having to track performance in a moment-by-moment fashion, such that any accumulating errors are shown and available for correction at the next glance of the display as the pilot's workload permits. In addition it allows for pilots to quickly assess the speed adjustment requirement by noting the relative amount and direction of the current speed indicator and the commanded speed indicator. Thus, the error-nulling algorithm/display demonstrated multiple advantages without any of the disadvantages of the other displays: (a) Its error-nulling capabilities led to excellent RTA performance; (b) it can be used strategically, as it is designed to provide the needed speed input information when the pilot's workload permits; and (c) it enables safe operation because the pilots can choose when to monitor the display without loss of overall RTA performance and without sacrificing out-the-window monitoring of traffic (i.e., enables efficient multitasking).

Although out of the scope of these experiments because they were limited to PFDs, future efforts might consider the use of HUD solutions, similar to that of Williams et al. (2006). However, as previous research has found, moving information from the flight deck avionics does eliminate eyes-in visual time and scanning time, but does not necessarily mitigate attentional issues and task allocation issues associated with using the information (see Fadden, Wickens, & Ververs, 2000, for a meta-analysis; and Prinzel & Risser, 2004, for a review of the issue of HUDs and attention). When the commanded speed was presented on a simulated HUD as in Williams et al. (2006), relatively poor RTA performance was still found. One would reasonably expect that the general findings in these experiments would still be obtained if the RTA speed information was moved to a HUD: Requiring pilots to track commanded speed on a moment-by-moment basis, as in Experiments 1 and 2, would still require increased attention and task resource allocation, whereas having to follow a commanded speed driven by the error-nulling algorithm (Experiment 3) would allow for more strategic task resource allocation and usage.

STM System Implications

The results of these experiments have 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 kt), and late with faster required speeds (i.e., 22 kt). This implies that ATC/STM algorithms should operate with middle-range speeds, similar to those associated with non-STBO taxi performance (see Hooey, Foyle, & Andre, 2000). 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 might be mitigated by imposing a small number of time constraint points each with its own RTAs. This has the resultant effect of turning a long route into a series of shorter routes—and thus improving RTA performance. An STBO Concept of Operations (see Hooey, Cheng, & Foyle, 2014) with a small number of intermediate time-constraint points along the taxi route and one at the departing runway, however, is feasible and might be safer, but only with flight-deck equipage similar to that used in Experiment 3 (i.e., 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) for his help with the conduct of Experiments 1 and 3, and to Dr. Kevin Jordan (San Jose State University) for his data analysis expertise. Portions of these experiments were previously presented at the 2009 Integrated Communication, Navigation and Surveillance, 2011 International Symposium on Aviation Psychology, and 2011 USA/Europe Air Traffic Management R&D Seminar conferences, and described in a San Jose State University master's thesis (Christina L. Kunkle). The authors thank the editors and reviewers for their helpful comments.

FUNDING

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

REFERENCES

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach*. Hillsdale, NJ: Erlbaum.

Cheng, V. H. L. (2002). Collaborative automation systems for enhancing airport surface traffic efficiency and safety. In *Proceedings of the 21st IEEE/AIAA Digital Avionics Systems Conference, Vol.1* (1D4-1–1D4-12). New York: Institute of Electrical and Electronics Engineers. doi:10.1109/DASC.2002.1067901

Cheng, V. H. L., Sharma, V., & Foyle, D. C. (2001). A study of aircraft taxi performance for enhancing airport surface traffic control. *IEEE Transactions on Intelligent Transportation Systems*, *2*, 39–54. doi:10.1109/6979.928715

Cheng, V. H. L., Yeh, A., Diaz, G. M., & Foyle, D. C. (2004). Surface-operation benefits of a collaborative automation concept. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference* (Paper AIAA 2004-5409). Reston, VA: American Institute of Aeronautics and Astronautics. doi:10.2514/6.2004-5409

EMMA2 Consortium. (2008). *A-SMGCS services, procedures, and operational requirements (SPOR), IP of 6th FP of European Commission/DG TREN*. Retrieved from www.dlr.de/emma2

Fadden, S., Wickens, C. D., & Ververs, P. M. (2000). Costs and benefits of head-up displays: An attention perspective and a meta analysis (Paper No. 2000-01-5542). In *2000 World Aviation Congress*. Warrendale, PA: Society of Automotive Engineers.

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 Ames Research Center. Retrieved from http://human-factors.arc.nasa.gov/groups/HCSL/publications/STBO%20ConOps_TM_2014_218354.pdf

Hooey, B. L., & Foyle, D. C. (2006). Pilot navigation errors on the airport surface: Identifying contributing factors and mitigating solutions. *The International Journal of Aviation Psychology*, *16*, 51–76. doi:10.1207/s15327108ijap1601_3

Hooey, B. L., Foyle, D. C., & Andre, A. D. (2000). Integration of cockpit displays for surface operations: The final stage of a human-centered design approach. *SAE Transactions: Journal of Aerospace*, *109*, 1053–1065.

Janssen, C. P., & Brumby, D. P. (2010). Strategic adaptation to performance objectives in a dual-task setting. *Cognitive Science*, *34*, 1548–1560. doi:10.1111/cogs.2010.34.issue-8

Joint Planning and Development Office. (2010). *Concept of operations for the Next Generation Air Transport System, version 3.2*. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a535795.pdf

Jones, D. R., Prinzel, L. J., Bailey, R. E., Arthur, J. J., & Barnes, J. R. (2014). Safely conducting airport surface trajectory-based operations. In *Proceedings of the 2014 IEEE/AIAA 33rd Digital Avionics Systems Conference (DASC) (7C2-1–7C2-16)*. New York: Institute of Electrical and Electronics Engineers. doi:10.1109/DASC.2014.6979518

National Academies Press (2015). *A review of the Next Generation Air Transportation System: Implications and importance of system architecture*. Retrieved from http://www.nap.edu/download.php?record_id=21721

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64. doi:10.1016/0010-0285(75)90004-3

Prinzel, L. J., III, & Risser, M. (2004). Head-up displays and attention capture (NASA/TM-2004-213000). Retrieved from http://ntrs.nasa.gov/search.jsp?R=20040065771

Rathinam, S., Montoya, J., & Jung, Y. (2008). An optimization model for reducing aircraft taxi times at the Dallas Fort Worth International Airport. In I. Grant (Ed.), *Proceedings of the 26th International Congress of the Aeronautical Sciences*. Retrieved from http://www.icas.org/icas_archive/icas2008/papers/570.pdf

SESAR Consortium. (2008). *SESAR master plan*. Retrieved from www.atmmasterplan.eu

Theunissen, E., Rademaker, R. M., Jinkins, R. D., & Uijt De Haag, M. (2002). Design and evaluation of taxi navigation displays In *Proceedings of the 21st Digital Avionics Systems Conference, Vol. 2* (11A3-1–11A3-13). New York: Institute of Electrical and Electronics Engineers. doi:10.1109/DASC.2002.1052969

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York, NY: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177. doi:10.1080/14639220210123806

Williams, J. L., Hooey, B. L., & Foyle, D. C. (2006). 4-D taxi clearances: Pilots' usage of time- and speed-based formats. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference* (Paper AIAA-2006-6611). Reston, VA: American Institute of Aeronautics and Astronautics. doi:10.2514/6.2006-6611